

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 06:37:13 ; Search time 15 Seconds

(without alignments)  
64.090 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTLCFASDAK 10

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	211	2 S25938	env protein - huma
2	52	100.0	219	2 S25939	env protein - huma
3	52	100.0	495	2 S31493	env polypeptide -
4	52	100.0	506	2 A40218	envelop glycoprote
5	52	100.0	729	1 VCLJJK	env polypeptide pr
6	52	100.0	843	1 H44001	env polypeptide pr
7	52	100.0	846	1 VCLJND	env polypeptide pr
8	52	100.0	847	2 T09448	envelope glycoprot
9	52	100.0	847	2 S13289	env protein - huma
10	52	100.0	851	2 S33985	env polypeptide -
11	52	100.0	852	1 T12016	envelope glycoprot
12	52	100.0	852	1 VCLJBR	envelope glycoprote
13	52	100.0	853	2 S54384	envelope polypept
14	52	100.0	854	2 S13288	env protein - huma
15	52	100.0	855	1 VCLJZR	env polypeptide pr
16	52	100.0	856	1 VCLJHR	env polypeptide pr
17	52	100.0	856	1 VCLJVL	env polypeptide pr
18	52	100.0	856	1 VCLJ3W	env polypeptide pr
19	52	100.0	856	1 A44963	env polypeptide pr
20	52	100.0	859	1 VCLJMN	env polypeptide pr
21	52	100.0	859	2 T01672	envelope polypept
22	52	100.0	861	1 VCLJIV	env polypeptide pr
23	52	100.0	861	1 VCLJKB	env polypeptide pr
24	52	100.0	861	1 VCLJSC	env polypeptide pr
25	52	100.0	868	1 VCLJH4	env polypeptide pr
26	49	94.2	855	1 VCLJAJ2	env polypeptide pr
27	42	80.8	854	1 VCLJSI	env polypeptide pr
28	42	80.8	877	1 S49197	envelope protein p
29	40	76.9	1476	2 A41185	alpha-2 macroglobu

30	39	75.0	609	2 A41081	alpha-1-inhibitor
31	39	75.0	1451	2 B41185	alpha-2 macroglobu
32	37	71.2	863	2 A53034	gag polypeptide -
33	36	69.2	329	2 S46807	hypothetical prote
34	36	69.2	514	2 A86547	hypothetical prote
35	36	69.2	514	2 E72076	polymorphic membra
36	36	69.2	554	2 E97852	hypothetical prote
37	36	69.2	555	2 A80321	glutamine-tRNA lig
38	35	67.3	350	2 E87327	hypothetical prote
39	35	67.3	515	2 F70786	probable pepa - My
40	35	67.3	554	1 SYEQOT	glutamine-tRNA lig
41	35	67.3	554	2 F90717	glutamine-tRNA syn
42	35	67.3	554	2 F85567	glutamine-tRNA syn
43	35	67.3	555	2 AC0585	glutamine-tRNA sy
44	34	65.4	287	2 C72099	conserved hypotnet
45	34	65.4	287	2 A86523	CT143 hypothetical

#### ALIGNMENTS

##### RESULT 1

S25938 env protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 25-Feb-1994 #sequence\_revision 30-Jan-1998 #text\_change 26-Aug-1999

C:Accession: S25938

R:Guo, H.G.; Chermann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher, H.

Nature 349, 745-746, 1991

A:Title: Sequence analysis of original HIV-1.

A:Reference number: S25937; MUID:51156044; PMID:2000145

A:Accession: S25938

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-211 <GUO>

A:Cross-references: EMBL:X57446; NID:960204; PIDN:CAA40692.1; PID:960205

A:Experimental source: strain JIB

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991

C:Genetics:

A:Gene: env

C:Superfamily: type B retrovirus env polypeptide

C:Keywords: coat protein

Query Match 100.0%; Score 52; DB 2; Length 211;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10

DB 9 TTLCFASDAK 18

##### RESULT 2

S25939 env protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 25-Feb-1994 #sequence\_revision 30-Jan-1998 #text\_change 26-Aug-1999

C:Accession: S25939

R:Guo, H.G.; Chermann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher, H.

Nature 349, 745-746, 1991

A:Title: Sequence analysis of original HIV-1.

A:Reference number: S25937; MUID:51156044; PMID:2000145

A:Accession: S25939

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-219 <GUO>

A:Cross-references: EMBL:X57447; NID:960202; PIDN:CAA40693.1; PID:9388536

A:Experimental source: strain JIB

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991

C:Genetics:

A:Gene: env

C:Superfamily: type B retrovirus env polypeptide

C:Keywords: coat protein

Query Match 100.0%; Score 52; DB 2; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10  
 |||||  
 Db 9 TTLCFASDAK 18

## RESULT 3

env polyprotein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 26-Aug-1999  
 C:Accession: S31493  
 R:Clc99, C.S.  
 submitted to the EMBL Data Library, December 1992  
 A:Reference number: S31493  
 A:Accession: S31493  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <CLB>  
 A:Cross-references: EMBL:Z19533; NID:G60244; PIDN:CAA79593.1; PID:G60245  
 C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: polyprotein

Query Match 100.0%; Score 52; DB 2; Length 495;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10  
 |||||  
 Db 23 TTLCFASDAK 32

## RESULT 4

env polyprotein gp120 - human immunodeficiency virus type 1  
 A:Accession: A40218  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 12-Apr-1995  
 C:Accession: A40218  
 R:Turner, S.; Tizard, R.; DeMarinis, J.; Pepinsky, R.B.; Zullo, J.; Schooley, R.; Fisher, Proc. Natl. Acad. Sci. U.S.A. 89, 1335-1339, 1992  
 A:Title: Resistance of primary isolates of human immunodeficiency virus type 1 to neutralization by monoclonal antibodies  
 A:Reference number: A40218; MUID:92159044; PMID:1741386  
 A:Accession: A40218  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-506 <TUR>  
 A:Experimental source: isolate P17  
 A>Note: sequence extracted from NCBI backbone (NCBI:P17:82240)  
 C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: glycoprotein

Query Match 100.0%; Score 52; DB 2; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10  
 |||||  
 Db 49 TTLCFASDAK 58

## RESULT 5

env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)  
 N:Alternate names: coat polyprotein  
 N:Contains: coat protein gp120; coat protein gp32  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A>Note: host Homo sapiens (man)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 01-Mar-1996  
 C:Accession: B42995

R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.  
 Virology 189, 534-546, 1992  
 A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated  
 A:Reference number: A42995; MUID:92351552; PMID:1323587

A:Accession: B42995  
 A:Molecule type: mRNA  
 A:Residues: 1-729 <SHI>  
 A:Cross-references: GB:S41266; GB:D01206  
 C:Genetics:

C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
 F:1-689/Domain: extracellular #status predicted <EXT>  
 F:1-729/Domain: signal sequence #status predicted <SIG>  
 F:17-33/Region: hydrophobic #status predicted  
 F:34-517/Product: coat protein gp120 #status predicted <CP1>  
 F:514-517/Region: cleavage processing #status predicted  
 F:518-729/Product: coat protein gp32 #status predicted <CP2>  
 F:518-534/Region: hydrophobic #status predicted  
 F:690-711/Domain: transmembrane #status predicted <TM1>  
 F:712-729/Domain: intracellular #status predicted <INT>  
 F:93,141,145,146,163,191,192,237,241,248,263,283,296,308,338,345,361,367,397,403,408,414

Query Match 100.0%; Score 52; DB 1; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 0.044;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10  
 |||||  
 Db 55 TTLCFASDAK 64

## RESULT 6

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)  
 N:Alternate names: coat polyprotein  
 N:Contains: coat protein gp120; coat protein gp41  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A>Note: host Homo sapiens (man)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Sep-1994  
 C:Accession: H44001  
 R:Li, Y.; Hui, H.; Burgees, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
 J. Virol. 66, 6587-6600, 1992  
 A:Title: Complete nucleotide sequence, genome organization, and biological properties of  
 A:Reference number: H44001; MUID:93021387; PMID:1404605  
 A:Accession: H44001  
 A:Molecule type: DNA  
 A:Residues: 1-843 <LTY>  
 A:Cross-references: GB:M93258  
 C:Genetics:

C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
 F:1-29/Domain: signal sequence #status predicted <SIG>  
 F:19-35/Region: hydrophobic  
 F:30-489/Product: coat protein gp120 #status predicted <GP1>  
 F:490-843/Product: coat protein gp41 #status predicted <GP2>  
 F:499-515/Region: hydrophobic  
 F:673-689/Region: hydrophobic  
 F:738-755/Domain: transmembrane #status predicted <TMN>  
 F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 100.0%; Score 52; DB 1; Length 843;  
 Best Local Similarity 100.0%; Pred. No. 0.05;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFASDAK 10  
 |||||  
 Db 49 TTLCFASDAK 58

## RESULT 7

VCLJND

env polyprotein precursor - human immunodeficiency virus type 1 (isolate NDK)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Accession: J00066  
R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;  
Gene 81, 275-284, 1989  
A:Title: Nucleotide sequence of HIV-1-NDK, a highly cytopathic strain of the human immunodeficiency virus  
A:Reference number: J00065; MUID:90034200; PMID:2806917  
A:Accession: J00066  
A:Molecule type: DNA  
A:Residues: 1-846 <SP1>  
A:Cross-references: GB:M27323; NID:G328154; PIDN:AAA44873.1; PID:G328162  
A:Note: the authors translated the codon GCG for residue 523 as Arg  
A:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-501/Product: coat protein gp120 #status predicted <CP1>  
F:502-846/Product: coat protein gp41 #status predicted <CP2>  
F:502-530/Domain: transmembrane #status predicted <TM1>  
F:674-692/Domain: transmembrane #status predicted <TM2>  
F:87,129,151,179,182,229,236,257,271,284,290,351,382,392,395,401,438,451,452,601,606

Query Match 100.0%; Score 52; DB 1; Length 846;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
|||||  
Db 49 TTLCASDAK 58

RESULT 8  
T09448  
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRF1)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
C:Accession: T09448  
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,  
submitted to the EMBL Data Library, July 1996  
A:Reference number: Z16673  
A:Accession: T09448  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-847 <PAN>  
A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 52; DB 2; Length 847;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
|||||  
Db 49 TTLCASDAK 58

RESULT 9  
S13289  
env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: S13289  
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A  
Nature 348, 69-73, 1990  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
A:Reference number: S13288; MUID:91043044; PMID:2172833

A:Accession: S13289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-847 <OBR>  
C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 52; DB 2; Length 847;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
|||||  
Db 49 TTLCASDAK 58

RESULT 10  
S33985  
env polyprotein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999  
C:Accession: S33985  
R:Carlini, F.  
submitted to the EMBL Data Library, November 1991  
A:Reference number: S33979  
A:Accession: S33985  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-851 <CAR>  
A:Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA77628.1; PID:G60199  
C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 52; DB 2; Length 851;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
|||||  
Db 50 TTLCASDAK 59

RESULT 11  
VCLUBR  
env polyprotein - human immunodeficiency virus type 1 (isolate BR)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Feb-1997  
C:Accession: A31667  
R:Amund, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S  
Virology 168, 79-89, 1989  
A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)  
A:Reference number: A94389; MUID:89085613; PMID:2789516  
A:Accession: A31667  
A:Molecule type: DNA  
A:Residues: 1-852 <ANA>  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; polyprotein; transmembrane protein  
F:1-516/Product: coat protein gp120 #status predicted <CP1>  
F:517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 100.0%; Score 52; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
|||||  
Db 50 TTLCASDAK 59

RESULT 12  
T12016  
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)

```

C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Shepard, W.H
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U90934; NID:G2351783; PIDN:AA059271.1; PID:G2351784
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide

Query Match          100.0%; Score 52; DB 2; Length 852;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
    |||||
Db 49 TTLCASDAK 58

RESULT 13
S54384
envelope polypeptide - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C:Accession: S54384
R:Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54384
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-853 <THE>
A:Cross-references: EMBL:M22639; NID:G329377; PIDN:AAA45370.1; PID:G329385
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: polypeptide

Query Match          100.0%; Score 52; DB 2; Length 853;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
    |||||
Db 49 TTLCASDAK 58

RESULT 14
S13288
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13288
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13288
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-854 <OBR>
C:Superfamily: type E retrovirus env polypeptide

Query Match          100.0%; Score 52; DB 2; Length 854;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
    |||||

```

```

Db 50 TTLCASDAK 59

RESULT 15
VCLJ3R
env polypeptide precursor - human immunodeficiency virus Zr-6
N:Alternate names: coat polypeptide
C:Species: human immunodeficiency virus Zr-6
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu
Gene 52, 71-82, 1987
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot
A:Reference number: AC6192; MUID:87248097; PMID:3036660
A:Accession: D26192
A:Molecule type: DNA
A:Residues: 1-855 <SRI>
A:Cross-references: GB:K03458; GB:M16322; NID:G129308; PIDN:AAA45380.1; PID:G329403
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypept
F:1-19/Domain: signal sequence #status predicted <SIG>
F:120-855/Product: env polypeptide #status predicted <MAT>
F:501-855/Product: transmembrane glycoprotein #status predicted <TMM>
F:87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,404

Query Match          100.0%; Score 52; DB 1; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
    |||||
Db 49 TTLCASDAK 58

Search completed: April 14, 2003, 06:39:01
Job time : 21 secs

```



FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 421 AA, 47493 MM, 25A575719C22967B CRC64;

Query Match 100.0%; Score 52; DB 1; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 0.0041;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
 Db 49 TTLCASDAK 58

RESULT 2  
 ENV\_HV123 STANDARD; PRT; 460 AA.  
 ID ENV\_HV123  
 AC P12491;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (Zaire 3 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11680;  
 RN [1] \_\_\_\_\_  
 RP MEDLINE=66259728; PubMed=3014529;  
 RA Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T., Buckler C.B., Martin M.A.;  
 RT "Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus."; Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).  
 RL -1- MISCELLANEOUS: THOUGH THIS SEQUENCE CONTAINS A COMPLETE ENV CODING REGION, INSERTION OF AN EXTRA NUCLEOTIDE CREATES A STOP CODON PRIOR TO THE NORMAL TERMINATION; THE AUTHORS SUGGEST THAT THIS VARIATION CAN ACCOUNT FOR THE LACK OF INFECTIVITY OF THIS CLONE.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL\_K03347; AAA45372.1;  
 DR EMBL\_K03347; ENVS23;  
 DR HIV\_K03347; ENVS23;  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.  
 KM SIGNAL.  
 FT CHAIN 1 29 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 202 BY SIMILARITY.  
 FT DISULFID 125 193 BY SIMILARITY.  
 FT DISULFID 130 149 BY SIMILARITY.  
 FT DISULFID 215 244 BY SIMILARITY.

FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 293 326 BY SIMILARITY.  
 FT DISULFID 372 439 BY SIMILARITY.  
 FT DISULFID 379 412 BY SIMILARITY.  
 FT DISULFID 397 404 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 460 AA, 51297 MM, 27B97EB75C7E9F50 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 0.0045;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
 Db 49 TTLCASDAK 58

RESULT 3  
 ENV\_HV1Y2 STANDARD; PRT; 843 AA.  
 ID ENV\_HV1Y2  
 AC P35961;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36377;  
 RN [1] \_\_\_\_\_  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93021387; PubMed=1404605;  
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H., Shaw G.M.;  
 RT "Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation." J. Virol. 66:6587-6600(1992).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL\_M93258; -; NOT\_ANNOTATED\_CDS.  
 DR PIR\_H44001; H44001.  
 DR InterPro: IPR000328; Env\_GP41.

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;
  Signal.
FT SIGNAL 1 29
FT CHAIN 30 489
FT TRANSMEM 490 843
FT TRANSMEM 738 755
FT DISULFID 53 73
FT DISULFID 118 201
FT DISULFID 125 192
FT DISULFID 130 155
FT DISULFID 214 243
FT DISULFID 224 235
FT DISULFID 292 326
FT DISULFID 373 432
FT DISULFID 380 405
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 135 135
FT CARBOHYD 138 138
FT CARBOHYD 154 154
FT CARBOHYD 158 158
FT CARBOHYD 184 184
FT CARBOHYD 193 193
FT CARBOHYD 230 230
FT CARBOHYD 237 237
FT CARBOHYD 258 258
FT CARBOHYD 272 272
FT CARBOHYD 285 285
FT CARBOHYD 291 291
FT CARBOHYD 297 297
FT CARBOHYD 327 327
FT CARBOHYD 351 351
FT CARBOHYD 381 381
FT CARBOHYD 389 389
FT CARBOHYD 395 395
FT CARBOHYD 400 400
FT CARBOHYD 400 400
FT CARBOHYD 435 435
FT CARBOHYD 450 450
FT CARBOHYD 598 598
FT CARBOHYD 603 603
FT CARBOHYD 612 612
FT CARBOHYD 624 624
FT CARBOHYD 803 803
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 843;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTFPCASDAK 10
    |||
Db 49 TTFPCASDAK 58

RESULT 4
ENV_HVIND STANDARD; PRT; 846 AA.
AC P18799;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=90034200; PubMed=2806917;

```

```

RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Gilbert F.,
RA Hampe A., Chermann J.C.,
RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus."
RL Gene 81:275-284(1989).
CC 1- MISCELLANEOUS: NDK. ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS. AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL; M27323; AAA44873.1; -.
DR PIR; J00066; VCLJND.
DR HIV; M27323; ENV5NDK.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;
  Signal.
FT SIGNAL 1 29
FT CHAIN 30 501
FT TRANSMEM 490 846
FT DISULFID 53 73
FT DISULFID 118 200
FT DISULFID 125 191
FT DISULFID 130 152
FT DISULFID 213 242
FT DISULFID 223 234
FT DISULFID 291 328
FT DISULFID 374 435
FT DISULFID 381 408
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 151 151
FT CARBOHYD 179 179
FT CARBOHYD 182 182
FT CARBOHYD 229 229
FT CARBOHYD 236 236
FT CARBOHYD 257 257
FT CARBOHYD 271 271
FT CARBOHYD 284 284
FT CARBOHYD 290 290
FT CARBOHYD 351 351
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 395 395
FT CARBOHYD 401 401
FT CARBOHYD 438 438
FT CARBOHYD 451 451
FT CARBOHYD 452 452
FT CARBOHYD 601 601
FT CARBOHYD 606 606
FT CARBOHYD 615 615
FT CARBOHYD 627 627
SQ SEQUENCE 846 AA; 96476 MW; 8A3B9DA527DE2E83 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 846;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTFPCASDAK 10
    |||
Db 49 TTFPCASDAK 58

RESULT 5

```



```

FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96466 MW; CD1B3D73AA5BC6E CRC64;

Query Match 100.0%; Score 52; DB 1; Length 847;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLCASDAK 10
Db 49 TTLCASDAK 58

RESULT 7
ENV_HV1UR STANDARD; PRT; 848 AA.
ID ENV_HV1UR
AC P20871;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11688;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38429; AAB03749.1; -
DR HIV; M38429; ENVSURCSF.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW signal.
FT SIGNAL 1 32
FT CHAIN 33 503 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 504 848 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.

```

```

FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 437 BY SIMILARITY.
FT DISULFID 381 410 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 848 AA; 96475 MW; 20767F51227ECF3 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 848;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLCASDAK 10
Db 49 TTLCASDAK 58

RESULT 8
ENV_HV1B8 STANDARD; PRT; 851 AA.
ID ENV_HV1B8
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=5111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Graybe J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL: K02011; AAA44661.1;
DR HIV: K02011; ENVSBR8.
DR GlycoSiteDB: P04582;
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00516; GP120.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 30
FT CHAIN 31 506
FT CHAIN 507 851
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 440
FT DISULFID 385 413
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 401 401
FT CARBOHYD 443 443
FT CARBOHYD 458 458
FT CARBOHYD 606 606
FT CARBOHYD 611 611
FT CARBOHYD 620 620
FT CARBOHYD 632 632
FT CARBOHYD 669 669
FT CARBOHYD 745 745
FT CARBOHYD 811 811
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CPe64;

Query Match 100.0%; Score 52; DB 1; Length 851;
Best Local Similarity 100.0%; Pired. No. 0.0061;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTLCFASDAK 10
Db 50 TTLCFASDAK 59

RESULT 9
ENV_HV1BN STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (IBR isolate) (HIV-1).

```

```

OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; Pubmed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -1 MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL: M21098; AAA44231.1;
DR PIP: A31667; VCLJBR.
DR HIV: M21098; ENVSBRVA.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 30
FT CHAIN 31 507
FT CHAIN 508 852
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 155
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 330
FT DISULFID 376 439
FT DISULFID 383 412
FT CARBOHYD 49 49
FT CARBOHYD 88 88
FT CARBOHYD 135 135
FT CARBOHYD 138 138
FT CARBOHYD 154 154
FT CARBOHYD 158 158
FT CARBOHYD 197 197
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 331 331
FT CARBOHYD 354 354
FT CARBOHYD 360 360
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 396 396
FT CARBOHYD 400 400
FT CARBOHYD 442 442
FT CARBOHYD 456 456
FT CARBOHYD 607 607
FT CARBOHYD 612 612
FT CARBOHYD 621 621
FT CARBOHYD 633 633
FT CARBOHYD 670 670
FT CARBOHYD 812 812
SQ SEQUENCE 852 AA; 97203 MW; 2B8B66345DE9315F CRC64;

```

Query Match	Score	DB	Length
Best Local Similarity 100.0%	Pred. No. 0.0081		
Matches 10, Conservative 0, Mismatches 0, Indels 0, Gaps 0			
Qy 1	TTLFCASDAPK 10		
Db 50	TTLFCASDAPK 59		
RESULT 10			
ENV_HV1S3	STANDARD;	PRT;	852 AA.
ID	ENV_HV1S3		
AC	P19549;		
DT	01-FEB-1991 (Rel. 17, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].		
DE	ENV.		
GN	Human immunodeficiency virus type 1 (SF3 isolate) (HIV-1).		
OS	Viruses; Retrocid viruses; Retroviridae; Lentivirus.		
OK	NCBI_TaxID=11690;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=90317906; PubMed=2370688;		
RX	York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;		
RT	"Human immunodeficiency virus type 1 cellular host range, replication, and cytopathicity are linked to the envelope region of the viral genome."		
RL	J. Virol. 64:4016-4020(1990).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; M38427; AAA45067.1; -		
DR	HIV; M38427; ENV5SF33.		
DR	InterPro; IPR000328; Env GP41		
DR	InterPro; IPR000777; GP120.		
DR	Pfam; PF00516; GP120; 1.		
DR	Pfam; PF00517; GP41; 1.		
KM	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.		
KM	Signal.		
FT	1 31	BY SIMILARITY.	
FT	CHAIN 32 506	EXTERIOR MEMBRANE GLYCOPROTEIN.	
FT	CHAIN 507 852	TRANSMEMBRANE GLYCOPROTEIN.	
FT	DISULFID 53 73	BY SIMILARITY.	
FT	DISULFID 118 206	BY SIMILARITY.	
FT	DISULFID 125 197	BY SIMILARITY.	
FT	DISULFID 130 156	BY SIMILARITY.	
FT	DISULFID 219 248	BY SIMILARITY.	
FT	DISULFID 229 240	BY SIMILARITY.	
FT	DISULFID 297 439	BY SIMILARITY.	
FT	DISULFID 377 439	BY SIMILARITY.	
FT	DISULFID 384 412	BY SIMILARITY.	
FT	CARBOHYD 87 87	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 129 129	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 136 136	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 141 141	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 142 142	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 155 155	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 159 159	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 189 189	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 198 198	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 242 242	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 263 263	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 277 277	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 290 290	N-LINKED (GLCNAC. . .) (POTENTIAL).	

FT DISULFID 229 240 BY SIMILARITY.  
 FT DISULFID 297 330 BY SIMILARITY.  
 FT DISULFID 376 442 BY SIMILARITY.  
 FT DISULFID 383 416 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 853 AA; 96721 MW; P9CD864DAAD07A5 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 853;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCFASDAK 10  
 DB 49 TTLCFASDAK 58

RESULT 12  
 ID ENV\_HVIMF STANDARD; PRT; 853 AA.  
 AC P19551;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11704;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317877; PubMed=1695254;  
 RA Stevenson M., Haggerty S., Lamontica C., Mann A.M., Meier C., Wasiaak A.;  
 RT "Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytopaths";  
 RL J. Virol. 64:3792-3803(1990).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL curation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M33943; AAA44850.1; -;  
 DR HIV: M33943; ENVSMFA.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW Aids; Coat protein; Polypeptide; glycoprotein; Transmembrane; Signal.  
 FT CHAIN 1 30  
 FT CHAIN 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 510 853 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 203 BY SIMILARITY.  
 FT DISULFID 126 194 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 216 245 BY SIMILARITY.  
 FT DISULFID 226 237 BY SIMILARITY.  
 FT DISULFID 294 443 BY SIMILARITY.  
 FT DISULFID 376 416 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 853 AA; 96912 MW; 3377B93B6F222BA CRC64;  
 Query Match 100.0%; Score 52; DB 1; Length 853;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCFASDAK 10  
 DB 50 TTLCFASDAK 59

RESULT 13  
 ID ENV\_HV122 STANDARD; PRT; 853 AA  
 AC P12487;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).

[illegible]

Oy	1	TTLFCASDAK 10		
Db	49	TTLFCASDAK 58		
 RESULT 14				
ENV_HV126	ID	_ENV_HV126	STANDARD;	PRT; 855 AA.
AC	P04580;			
DC	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120), Transmembrane glycoprotein (GP41)].			
CN	ENV.			
OS	Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxId=11708;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87248097; PubMed=3036660;			
RA	Srinivasan A., Anand R., York D., Ranganathan P., Feorino P., Schiochetman G., Curran J., Kalayanaram V.S., Luciw P.A., Sanchez-Pescador R.;			
RA	Zaire; Molecular characterization of human immunodeficiency virus from Zaire; nucleotide sequence analysis identifies conserved and variable domains in the envelope gene.";			
RL	Gene 52:71-82(1987).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; K03458; AAA5380.1; -.			
DR	PIR; D26192; YCLJ2R.			
DR	HIV; K03458; ENV526.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00517; GP41; 1.			
DR	pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	510	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	511	855	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISUPLD	53	73	BY SIMILARITY.
FT	DISUPLD	118	207	BY SIMILARITY.
FT	DISUPLD	125	198	BY SIMILARITY.
FT	DISUPLD	130	155	BY SIMILARITY.
FT	DISUPLD	220	249	BY SIMILARITY.
FT	DISUPLD	230	241	BY SIMILARITY.
FT	DISUPLD	298	332	BY SIMILARITY.
FT	DISUPLD	378	444	BY SIMILARITY.
FT	DISUPLD	385	417	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	129	129	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	145	145	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	189	189	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	236	236	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	243	243	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	264	264	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	278	278	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	291	291	N-LINKED (GLCNAC. . ) (POTENTIAL).
FT	CARBOHYD	297	297	N-LINKED (GLCNAC. . ) (POTENTIAL).

```

FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
DB 49 TTLCASDAK 50

RESULT 15
ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375,
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11678;
RN [1]
RX MEDLINE=8511123; PubMed=2578615;
RA Patner L., Haseltine W., Patarca P., Liyak F.J., Starcich B.P.,
RA Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Fagaz T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.",
RT Nature 313:277-284 (1985).
RL [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells.",
RL J. Biol. Chem. 268:10373-10382 (1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: M15654; AAA44205.1;
DR PIR: A03973; VCLJH3
DR HIV: M15654; ENVSBH102.

```

```

DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT CHAIN 1 30
FT CHAIN 512 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88
FT CARBOHYD 135 135
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 625 625
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18911BB27 CRC64;

Query Match 100.0%; Score 52; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10
DB 50 TTLCASDAK 59

```

Search completed: April 14, 2003, 06:38:33  
Job time : 27 secs



```

ID 09DU21      PRELIMINARY;      PRT;      64 AA.
AC 09DU21;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Env protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIDU15;
RX MEDLINE=21002573; PubMed=11118069;
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Lwin H.H., Tun K.M.,
   Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;
RT "Emergence of new forms of human immunodeficiency virus type 1
   intersubtype recombinants in central myanmar.";
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).
DR EMBL: AB043903; BAB19234.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER
SQ SEQUENCE 64 AA; 7325 MW; 07815947877D3B33 CRC64;

```

```

Query Match      100.0%; Score 52; DB 15; Length 64;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTLCASDAK 10
DB 49 TTLCASDAK 58

```

```

RESULT 3
09DU09          PRELIMINARY;      PRT;      66 AA.
ID 09DU09;
AC 09DU09;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Env protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCSW2;
RX MEDLINE=21002573; PubMed=11118069;
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Lwin H.H., Tun K.M.,
   Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;
RT "Emergence of new forms of human immunodeficiency virus type 1
   intersubtype recombinants in central myanmar.";
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).
DR EMBL: AB043905; BAB19246.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER
SQ SEQUENCE 66 AA; 7661 MW; 05952A0805568349 CRC64;

```

```

Query Match      100.0%; Score 52; DB 15; Length 66;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTLCASDAK 10
DB 49 TTLCASDAK 58

```

RESULT 4

```

09DU03          PRELIMINARY;      PRT;      67 AA.
ID 09DU03;
AC 09DU03;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Env protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCSW12;
RX MEDLINE=21002573; PubMed=11118069;
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Lwin H.H., Tun K.M.,
   Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;
RT "Emergence of new forms of human immunodeficiency virus type 1
   intersubtype recombinants in central myanmar.";
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).
DR EMBL: AB043906; BAB19252.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER
SQ SEQUENCE 67 AA; 7718 MW; 4274652A0805FD58 CRC64;

```

```

Query Match      100.0%; Score 52; DB 15; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTLCASDAK 10
DB 49 TTLCASDAK 58

```

```

RESULT 5
09OKJ9          PRELIMINARY;      PRT;      79 AA.
ID 09OKJ9;
AC 09OKJ9;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08106V4;
RX MEDLINE=99094949; PubMed=9878014;
RA Van Dyke R.B., Korber B.T., Popek E., Macken C., Widmayer S.M.,
   Bardegenet A., Hansen I.C., Wiznia A., Luzzitiga K., Viscarello R.R.,
   Molinsky S., the Ariel Core Investigators;
RT "The Ariel Project: A prospective cohort study of maternal-child
   transmission of human immunodeficiency virus type 1 in the era of
   maternal antiretroviral therapy.";
RL J. Infect. Dis. 179:319-328(1999).
DR EMBL: AF112539; AAF13317.1; -
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER
SQ SEQUENCE 79 AA; 8854 MW; 731EE519A785BAC CRC64;

```

```

Query Match      100.0%; Score 52; DB 15; Length 79;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTLCASDAK 10
DB 34 TTLCASDAK 43

```

## RESULT 6

Q9DU34 PRELIMINARY; PRT; 82 AA.  
AC Q9DU34;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Env protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MIDU10;  
RX MEDLINE=21002573; PubMed=1118069;  
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Lwin H.H., Tun K.M.,  
RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;  
RT "Emergence of new forms of human immunodeficiency virus type 1  
RT intersubtype recombinants in central Myanmar.";  
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).  
DR EMBL; AB043900; BAB19221.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KM AIDS; Coat protein; Glycoprotein.  
FT NON\_TER  
SQ SEQUENCE 82 AA; 9286 MW; 489E7B64117B18E CRC64;

Query Match 100.0%; Score 52; DB 15; Length 82;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 49 TTLCASDAK 58

## RESULT 7

Q9DU15 PRELIMINARY; PRT; 82 AA.  
AC Q9DU15;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Env protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MIDU18;  
RX MEDLINE=21002573; PubMed=1118069;  
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Lwin H.H., Tun K.M.,  
RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;  
RT "Emergence of new forms of human immunodeficiency virus type 1  
RT intersubtype recombinants in central Myanmar.";  
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).  
DR EMBL; AB043904; BAB19240.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KM AIDS; Coat protein; Glycoprotein.  
FT NON\_TER  
SQ SEQUENCE 82 AA; 9316 MW; EA31D130362BF090 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 82;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 11 TTLCASDAK 10

Db 49 TTLCASDAK 58

## RESULT 8

Q74305 PRELIMINARY; PRT; 83 AA.  
AC Q74305;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein gp120 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ray S.C., Lubaki N.M., Dhruva B.R., Siliciano R.F., Bollinger R.C.;  
RT "Strain-specific Cytolytic T lymphocyte responses directed against  
RT Human Immunodeficiency Virus type 1 env.";  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U78832; AAB02636.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KM AIDS; Coat protein; Glycoprotein.  
FT NON\_TER  
SQ SEQUENCE 83 AA; 9606 MW; 340C0B1F92918EC1 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 50 TTLCASDAK 59

## RESULT 9

Q8URC6 PRELIMINARY; PRT; 96 AA.  
AC Q8URC6;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98ZM066F;  
RX MEDLINE=21602569; PubMed=11739704;  
RA Trask S.A., Dedejyn C.A., Fidele U., Chen Y., Meleth S., Kasolo F.,  
RA Musonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
RT "Molecular epidemiology of human immunodeficiency virus type 1  
RT transmission in a heterosexual cohort of discordant couples in  
RT Zambia.";  
RL J. Virol. 76:397-405(2002).  
DR EMBL; AF405016; AAB66542.1; -  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
FT NON\_TER  
SQ SEQUENCE 96 AA; 11070 MW; 4AA9696D348B0D04 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 96;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 18 TTLCASDAK 27

## RESULT 10

Q8URC5 PRELIMINARY; PRT; 96 AA.  
 AC Q8URC5;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=98ZM066M;  
 RX MEDLINE=21602569; PubMed=11739704;  
 RA Traak S.A., Derdeyn C.A., Fidell U., Chen Y., Meleth S., Kasolo F.,  
 Musonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
 RT "Molecular epidemiology of human immunodeficiency virus type 1  
 transmission in a heterosexual cohort of discordant couples in  
 Zambia";  
 RL J. Virol. 76:397-405(2002).  
 DR EMBL, AF405017; AAL66543.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 FT NON\_TER 1  
 FT NON\_TER 96  
 SQ SEQUENCE 96 AA; 11070 MW; 4AA9696D309F5C45 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 0.0031;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TTLCASDAK 10  
 Db 18 TTLCASDAK 27

## RESULT 11

Q8URD8 PRELIMINARY; PRT; 97 AA.  
 AC Q8URD8;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=98ZM126F;  
 RX MEDLINE=21602569; PubMed=11739704;  
 RA Traak S.A., Derdeyn C.A., Fidell U., Chen Y., Meleth S., Kasolo F.,  
 Musonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
 RT "Molecular epidemiology of human immunodeficiency virus type 1  
 transmission in a heterosexual cohort of discordant couples in  
 Zambia";  
 RL J. Virol. 76:397-405(2002).  
 DR EMBL, AF405000; AAL66530.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 FT NON\_TER 1  
 FT NON\_TER 97  
 SQ SEQUENCE 97 AA; 11186 MW; 8837DEC434627102 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.0032;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
 Db 19 TTLCASDAK 28

## RESULT 12

Q8URD6 PRELIMINARY; PRT; 97 AA.  
 AC Q8URD6;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=98ZM075F;  
 RX MEDLINE=21602569; PubMed=11739704;  
 RA Traak S.A., Derdeyn C.A., Fidell U., Chen Y., Meleth S., Kasolo F.,  
 Musonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
 RT "Molecular epidemiology of human immunodeficiency virus type 1  
 transmission in a heterosexual cohort of discordant couples in  
 Zambia";  
 RL J. Virol. 76:397-405(2002).  
 DR EMBL, AF405002; AAL66532.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 FT NON\_TER 1  
 FT NON\_TER 97  
 SQ SEQUENCE 97 AA; 11216 MW; 86F585AD1939146A CRC64;

Query Match 100.0%; Score 52; DB 15; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.0032;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TTLCASDAK 10  
 Db 19 TTLCASDAK 28

## RESULT 13

Q8URD5 PRELIMINARY; PRT; 97 AA.  
 AC Q8URD5;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=98ZM075M;  
 RX MEDLINE=21602569; PubMed=11739704;  
 RA Traak S.A., Derdeyn C.A., Fidell U., Chen Y., Meleth S., Kasolo F.,  
 Musonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
 RT "Molecular epidemiology of human immunodeficiency virus type 1  
 transmission in a heterosexual cohort of discordant couples in  
 Zambia";  
 RL J. Virol. 76:397-405(2002).  
 DR EMBL, AF405003; AAL66533.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 FT NON\_TER 1  
 FT NON\_TER 97  
 SQ SEQUENCE 97 AA; 11239 MW; 86F585AD1939146A CRC64;

Query Match 100.0%; Score 52; DB 15; Length 97;

Best Local Similarity 100.0%; Pred.No.0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 19 TTLCASDAK 28

## RESULT 14

Q8URD2 PRELIMINARY; PRT; 97 AA.  
AC Q8URD2;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98ZM069F;  
RA Trask S.A., Derdeyn C.A., Fideli U., Chen Y., Meleth S., Kasolo F.,  
RA Musonda R., Hunter B., Gao F., Allen S., Hahn B.H.;  
RT "Molecular epidemiology of human immunodeficiency virus type 1  
RT transmission in a heterosexual cohort of discordant couples in  
RT Zambia.";  
RL J. Virol. 76:397-405 (2002).  
DR EMBL; AF405010; AAL6536.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
FT NON\_TER 1  
FT NON\_TER 97  
SQ SEQUENCE 97 AA; 11097 MW; F8427A8FDC26B5A7 CRC64;

Query Match 100.0%; Score 52; DB 15; Length 97;  
Best Local Similarity 100.0%; Pred.No.0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 19 TTLCASDAK 28

## RESULT 15

Q8URD1 PRELIMINARY; PRT; 97 AA.  
AC Q8URD1;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98ZM069M;  
RA Trask S.A., Derdeyn C.A., Fideli U., Chen Y., Meleth S., Kasolo F.,  
RA Musonda R., Hunter B., Gao F., Allen S., Hahn B.H.;  
RT "Molecular epidemiology of human immunodeficiency virus type 1  
RT transmission in a heterosexual cohort of discordant couples in  
RT Zambia.";  
RL J. Virol. 76:397-405 (2002).  
DR EMBL; AF405011; AAL6537.1; -.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
FT NON\_TER 1  
FT NON\_TER 97  
SQ SEQUENCE 97 AA; 11097 MW; F8427A8FDC26B5A7 CRC64;

SQ SEQUENCE 97 AA; 11141 MW; 36D730CAD1D9954A CRC64;

Query Match 100.0%; Score 52; DB 15; Length 97;  
Best Local Similarity 100.0%; Pred.No.0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 19 TTLCASDAK 28

Search completed: April 14, 2003, 06:40:53  
Job time : 86 secs

1



PT infection and cancer  
 XX  
 PS Claim 19; Page 39; 49pp; English.  
 XX  
 CC The present sequence represents a specific example of an immunogenic  
 CC peptide which was used in a new method of inducing a cytotoxic T cell  
 CC (CTL) response against a preselected antigen in a patient. The method  
 CC comprises contacting CTLs from the patient with the immunogenic peptide  
 CC (containing defined motifs) which binds one of the four HLA MHC products  
 CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation  
 CC constant (Kd) of less than 5 x 10<sup>-7</sup> M. Immunogens are viral, e.g. human  
 CC immunodeficiency virus type 1 (HIV-1), hepatitis B virus (HBV) and  
 CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for  
 CC the prevention and treatment of viral infection and cancer. The  
 CC immunogens may be administered to the patient as a nucleic acid encoding  
 CC the peptide (gene vaccine).  
 XX  
 SQ Sequence 10 AA;  
 Query Match 100.0%; Score 52; DB 18; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0068;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTLPFASDAK 10  
 Db 1 TTLPFASDAK 10  
 RESULT 2  
 AAY66220  
 ID AAY66220 standard; Peptide; 10 AA.  
 XX  
 AC AAY66220;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE HLA-A28-binding HIV-1 Env peptide #22.  
 XX  
 KM HIV-1; MHC; major histocompatibility complex; Class I; HLA;  
 KM human leukocyte antigen; allele; binding; conserved; genome;  
 KM peptide; targeting; toxic; drug; antibody; antigen; antiviral;  
 KM molecular conjugate; therapeutic; diagnosis; treatment; pathogen;  
 KM localisation; quantification; detection; infection; drug resistance;  
 KM immune response.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9949893-A1.  
 XX  
 PD 07-OCT-1999.  
 XX  
 PF 31-MAR-1999; 99WO-US07111.  
 XX  
 PR 31-MAR-1998; 98US-0052530.  
 XX  
 PA (UYBO-) UNIV BOSTON.  
 XX  
 PI Delisi C, Berzofsky J, Gulukota K, Vaccaro D, Weng Z, Zhang C;  
 XX  
 DR WPI; 2000-038361/03.  
 XX  
 PT Novel methods for designing molecular conjugate therapeutics which are  
 XX used for diagnosis, imaging and treatment against pathogens  
 XX  
 PS Example 3; Page 45; 62pp; English.  
 XX  
 CC AAY66199-Y66413 are peptides derived from conserved portions of the  
 CC HIV-1 genome which bind to different HLA alleles of MHC (major  
 CC histocompatibility complex) Class I molecules. The peptides are used to  
 CC construct targeting antigens comprising one or more peptides bound to  
 CC the corresponding MHC Class I molecule, which can be used to raise  
 CC antibodies. The antibody may then be used as a targeting vehicle to  
 CC deliver a potentially toxic drug to its target site of action, rather

CC than administering it systemically, which may result in adverse side  
 CC effects. The invention relates to improved methods for the design of  
 CC molecular conjugate therapeutics for the diagnosis and treatment of  
 CC infections caused by pathogens with a high mutation rate (such as  
 CC HIV-1). This method involves identifying conserved peptide-encoding  
 CC regions among the genomes of multiple variants of a pathogen, identifying  
 CC the Class I MHC molecules which occur with greatest frequency in a  
 CC population of interest (e.g., human sub-populations), and determining  
 CC which of the peptides bind to the Class I MHC molecules. The MHC-binding  
 CC peptides and the corresponding Class I MHC molecules are selected and  
 CC used to construct targeting antigens, which are in turn used to produce  
 CC targeting antibodies. The methods may be used in localisation,  
 CC quantification and in situ detection of specific peptide-MHC Class I  
 CC complexes and also to detect and treat viral infection. The methods of  
 CC the invention mitigate against the development of viral resistance to  
 CC drugs and to the immune response, as well as providing a solution for  
 CC targeting toxic compounds to destroy viruses sequestered in sites not  
 CC accessible to T cells. In addition, the methods eliminate the virus,  
 CC whereas current therapies only arrest viral replication.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 100.0%; Score 52; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0068;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTLPFASDAK 10  
 Db 1 TTLPFASDAK 10  
 RESULT 3  
 AAY73189  
 ID AAY73189 standard; Peptide; 10 AA.  
 XX  
 AC AAY73189;  
 XX  
 DT 28-FEB-2000 (first entry)  
 XX  
 DE HIV-derived MHC class I (CTL) epitope, #347.  
 XX  
 KM Chimeric; pan DR epitope; expression vector;  
 KM promoter; major histocompatibility complex; MHC; targeting; peptide;  
 KM epitope; antigen; presentation; class I; cytotoxic pathway;  
 KM endoplasmic reticulum; class II; extracellular antigen;  
 KM endocytic pathway; helper T lymphocyte; HTL; universal epitope;  
 KM cytotoxic T lymphocyte; CTL; immune response; immunogenicity; assay;  
 KM vaccine; immunity; infection; pathogen; virus; HIV, HBV, HCV,  
 KM hepatitis B; hepatitis C; bacterium; protozoan; tumour cell;  
 KM autoimmune disease; activation; antiviral; antimalarial;  
 KM immunoprotective.  
 XX  
 OS Synthetic.  
 XX  
 PN Human immunodeficiency virus.  
 XX  
 PD WO9958658-A2.  
 XX  
 PI 18-NOV-1999.  
 XX  
 PF 13-MAY-1999; 99WO-US10646.  
 XX  
 PR 13-MAY-1998; 98US-0078904.  
 XX  
 PR 15-MAY-1998; 98US-0085751.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Flies JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
 XX  
 PI Chesnut RW;  
 XX  
 DR WPI; 2000-039103/03.  
 XX  
 PT Expression vectors encoding major histocompatibility targeting  
 PT sequence, used as, e.g. tumor vaccines -

XX Claim 11; Page 71; 130pp; English.

PS  
XX  
CC Sequences AA73175-Y73266 represent human immunodeficiency virus (HIV)-  
CC derived MHC class I (CTL) epitopes which are claimed for use in the  
CC present invention. The invention relates to a novel expression vector  
CC comprising a promoter operably linked to a fusion gene encoding a major  
CC histocompatibility complex (MHC) targeting sequence, and two or more  
CC heterologous peptide epitopes. The MHC targeting sequence may be a  
CC class I targeting sequence, which directs an MHC class I epitope to  
CC a cytosolic pathway or to the endoplasmic reticulum, or an MHC class  
CC I targeting sequence, which directs extracellular antigens to  
CC enter the endocytic pathway to be processed into antigen peptides  
CC for presentation on MHC class II molecules. The heterologous  
CC epitopes may comprise either helper T lymphocyte (HTL) epitopes,  
CC or a cytotoxic T lymphocyte (CTL) epitope and a universal HTL  
CC epitope such as a pan DR epitope (PADRE). The vectors are useful  
CC for stimulating an immune response in vivo, as well as for use in  
CC assaying the human immunogenicity of a human T cell peptide epitope in  
CC vivo in a non-human mammal. They provide a nucleic acid vaccine for  
CC enhancing immunity against infectious pathogens, such as viruses (e.g.,  
CC HIV, hepatitis B (HBV) and hepatitis C (HCV)), bacteria, protozoa (e.g.,  
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
CC combined with other MHC class I and class II epitopes to increase the  
CC number of cells that are activated in response to a given antigen and  
CC provide a broader population coverage of MHC-reactive alleles.

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLPFCASDAK 10  
| | | | | | | | | |  
Db 1 TTLPFCASDAK 10

RESULT 4  
ABP14352  
ID ABP14352 standard; Peptide; 10 AA.  
AC  
XX ABP14352;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE HIV A03 super motif env peptide #92.  
XX  
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KW vpr; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;  
KW antigen; vaccine; HIV infection; immunisation; virucide.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200124810-A1.  
XX  
PD 12-APR-2001.  
XX  
PF 05-OCT-2000; 2000MO-US27766.  
XX  
PR 05-OCT-1999; 99US-0412863.  
XX  
PA (EPIM-) EPIMMUNE INC.  
XX  
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX  
DR WPI; 2001-354887/37.  
XX  
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
XX peptide groups, useful for vaccinating against HIV-1 -

PS Claim 32; Page 162; 448pp; English.

XX  
CC The present invention describes a composition (I) comprising a prepared  
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
CC ABL25397). (I) has virucide activity and can be used in vaccines. (I)  
CC may be used for immunising subjects against HIV-1 infections. The use of  
CC group-based vaccines has several advantages over traditional vaccines,  
CC particularly when compared to the use of whole antigens in vaccine  
CC compositions. There is evidence that the immune response to whole  
CC antigens is directed largely toward variable regions of the antigen,  
CC allowing for immune escape due to mutations. The groups for inclusion in  
CC an group-based vaccine may be selected from conserved regions of viral or  
CC tumour-associated antigens, which therefore reduces the likelihood of  
CC escape mutants. Furthermore, immunosuppressive groups that may be present  
CC in whole antigens can be avoided with the use of group-based vaccines.  
CC An additional advantage of an group-based vaccine approach is the ability  
CC to combine selected groups (CTL and HTL), and further, to modify the  
CC composition of the groups, achieving, for example, enhanced  
CC immunogenicity. Accordingly, the immune response can be modulated, as  
CC appropriate, for the target disease. Similar engineering of the response  
CC is not possible with traditional approaches. ABP1501 to ABP25412  
CC represent peptide sequences used in the exemplification of the present  
CC invention.

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLPFCASDAK 10  
| | | | | | | | | |  
Db 1 TTLPFCASDAK 10

RESULT 5  
ABP20295  
ID ABP20295 standard; Peptide; 10 AA.  
AC  
XX ABP20295;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE HIV A03 motif env peptide #499.  
XX  
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KW vpr; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;  
KW antigen; vaccine; HIV infection; immunisation; virucide.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200124810-A1.  
XX  
PD 12-APR-2001.  
XX  
PF 05-OCT-2000; 2000MO-US27766.  
XX  
PR 05-OCT-1999; 99US-0412863.  
XX  
PA (EPIM-) EPIMMUNE INC.  
XX  
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX  
DR WPI; 2001-354887/37.  
XX  
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
XX peptide groups, useful for vaccinating against HIV-1 -  
XX  
PS Claim 32; Page 285; 448pp; English.  
XX  
CC The present invention describes a composition (I) comprising a prepared

CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABL25397). (1) has virucide activity and can be used in vaccines. (1)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABL1501 to ABL25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.

XX Sequence 10 AA;  
 SQ

Query Match 100.0%; Score 52; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0068;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
 |||||  
 Db 1 TTLCASDAK 10

RESULT 6  
 ABL20315  
 ID ABL20315 standard; Peptide; 10 AA.  
 AC ABL20315;  
 XX  
 DT 15-JUL-2002 (first entry)  
 DE HIV A03 motif env peptide #519.  
 XX  
 KW HIV, HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 KW vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
 KW antigen; vaccine; HIV infection; immunisation; virucide.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200124810-A1.  
 XX  
 PD 12-APR-2001.  
 XX  
 PR 05-OCT-2000; 2000WO-US27766.  
 PF 05-OCT-1999; 99US-0412863.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX  
 DR WPI; 2001-354887/37.  
 XX  
 PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1 -  
 XX  
 PS Claim 32; Page 286; 448pp; English.  
 XX  
 CC The present invention describes a composition (1) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABL25397). (1) has virucide activity and can be used in vaccines. (1)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABL1501 to ABL25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.

CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABL1501 to ABL25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.

XX Sequence 10 AA;  
 SQ

Query Match 100.0%; Score 52; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0068;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
 |||||  
 Db 1 TTLCASDAK 10

RESULT 7  
 ABL22610  
 ID ABL22610 standard; Peptide; 10 AA.  
 AC ABL22610;  
 XX  
 DT 15-JUL-2002 (first entry)  
 DE HIV A11 motif env peptide #333.  
 XX  
 KW HIV, HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 KW vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
 KW antigen; vaccine; HIV infection; immunisation; virucide.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200124810-A1.  
 XX  
 PD 12-APR-2001.  
 XX  
 PR 05-OCT-2000; 2000WO-US27766.  
 PF 05-OCT-1999; 99US-0412863.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX  
 DR WPI; 2001-354887/37.  
 XX  
 PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1 -  
 XX  
 PS Claim 32; Page 332; 448pp; English.  
 XX  
 CC The present invention describes a composition (1) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABL25397). (1) has virucide activity and can be used in vaccines. (1)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABL1501 to ABL25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.

CC compositions. There is evidence that the immune response to whole  
CC antigens is directed largely toward variable regions of the antigen.  
CC allowing for immune escape due to mutations. The groups for inclusion in  
CC an group-based vaccine may be selected from conserved regions of viral or  
CC tumour-associated antigens, which therefore reduces the likelihood of  
CC escape mutants. Furthermore, immunosuppressive groups that may be present  
CC in whole antigens can be avoided with the use of group-based vaccines.  
CC An additional advantage of an group-based vaccine approach is the ability  
CC to combine selected groups (CTL and HTL), and further, to modify the  
CC composition of the groups, achieving, for example, enhanced  
CC immunogenicity. Accordingly, the immune response can be modulated, as  
CC appropriate, for the target disease. Similar engineering of the response  
CC is not possible with traditional approaches. ABP1501 to ABP25412  
CC represent peptide sequences used in the exemplification of the present  
CC invention.

SO Sequence 10 AA;

Query Match 100.0%; Score 52; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLFCASDAK 10  
| | | | | | | | | |  
Db 1 TTLFCASDAK 10

RESULT 8  
ABP25059  
ID ABP25059 standard; Peptide; 10 AA.  
AC ABP25059;  
XX 15-JUL-2002 (first entry)  
DT  
XX HIV A3 supermotif peptide #8.  
DE  
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KW vpr; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
KW antigen; vaccine; HIV infection; immunisation; virucide.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
XX WO200124810-A1.  
XX  
XX 12-APR-2001.  
PD  
XX 05-OCT-2000; 2000WO-US27766.  
PF  
XX 05-OCT-1999; 99US-0412863.  
PR  
XX (EPIM-) EPIMMUNE INC.  
PA  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX WPI; 2001-354887/37.  
DR  
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
PT peptide groups, useful for vaccinating against HIV-1 -  
PT  
XX Disclosure; Page 415; 448pp; English.  
XX  
XX The present invention describes a composition (I) comprising a prepared  
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
CC may be used for immunising subjects against HIV-1 infections. The use of  
CC group-based vaccines has several advantages over traditional vaccines,  
CC particularly when compared to the use of whole antigens in vaccine  
CC compositions. There is evidence that the immune response to whole  
CC antigens is directed largely toward variable regions of the antigen,  
CC allowing for immune escape due to mutations. The groups for inclusion in

CC an group-based vaccine may be selected from conserved regions of viral or  
CC tumour-associated antigens, which therefore reduces the likelihood of  
CC escape mutants. Furthermore, immunosuppressive groups that may be present  
CC in whole antigens can be avoided with the use of group-based vaccines.  
CC An additional advantage of an group-based vaccine approach is the ability  
CC to combine selected groups (CTL and HTL), and further, to modify the  
CC composition of the groups, achieving, for example, enhanced  
CC immunogenicity. Accordingly, the immune response can be modulated, as  
CC appropriate, for the target disease. Similar engineering of the response  
CC is not possible with traditional approaches. ABP1501 to ABP25412  
CC represent peptide sequences used in the exemplification of the present  
CC invention.

SO Sequence 10 AA;

Query Match 100.0%; Score 52; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLFCASDAK 10  
| | | | | | | | | |  
Db 1 TTLFCASDAK 10

RESULT 9  
ABP25176  
ID ABP25176 standard; Peptide; 10 AA.  
AC ABP25176;  
XX 15-JUL-2002 (first entry)  
DT  
XX HIV derived peptide #45.  
DE  
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KW vpr; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
KW antigen; vaccine; HIV infection; immunisation; virucide.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
XX WO200124810-A1.  
XX  
XX 12-APR-2001.  
PD  
XX 05-OCT-2000; 2000WO-US27766.  
PF  
XX 05-OCT-1999; 99US-0412863.  
PR  
XX (EPIM-) EPIMMUNE INC.  
PA  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX WPI; 2001-354887/37.  
DR  
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
PT peptide groups, useful for vaccinating against HIV-1 -  
PT  
XX Example 2; Page 422; 448pp; English.  
XX  
XX The present invention describes a composition (I) comprising a prepared  
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
CC may be used for immunising subjects against HIV-1 infections. The use of  
CC group-based vaccines has several advantages over traditional vaccines,  
CC particularly when compared to the use of whole antigens in vaccine  
CC compositions. There is evidence that the immune response to whole  
CC antigens is directed largely toward variable regions of the antigen,  
CC allowing for immune escape due to mutations. The groups for inclusion in  
CC an group-based vaccine may be selected from conserved regions of viral or  
CC tumour-associated antigens, which therefore reduces the likelihood of  
CC escape mutants. Furthermore, immunosuppressive groups that may be present

CC	in whole antigens can be avoided with the use of group-based vaccines.
CC	An additional advantage of an group-based vaccine approach is the ability
CC	to combine selected groups (CTL and HTL), and further, to modify the
CC	composition of the groups, achieving, for example, enhanced
CC	immunogenicity. Accordingly, the immune response can be modulated, as
CC	appropriate, for the target disease. Similar engineering of the response
CC	is not possible with traditional approaches. ABP11501 to ABP25412
CC	represent peptide sequences used in the exemplification of the present
CC	invention.
XX	
XX	Sequence 10 AA;
XX	
XX	Query Match 100.0%; Score 52; DB 22; Length 10;
XX	Best Local Similarity 100.0%; Pred. No. 0.0068;
XX	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TTTFCASDAK 10
QY	
QY	1 TTTFCASDAK 10
Db	
Db	1 TTTFCASDAK 10
RESULT 10	
ABP25246	
ID	ABP25246 standard; Peptide; 10 AA.
XX	
XX	ABP25246;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	HIV-derived A3-supertype peptide #18.
XX	
KW	HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vif; vpr;
KW	vpr; vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope;
KW	antigen; vaccine; HIV infection; immunisation; viraucide.
XX	
OS	Human immunodeficiency virus type 1.
XX	
XX	WO200124810-A1.
PN	
XX	12-APR-2001.
PD	
PF	05-OCT-2000; 2000WO-US27766.
XX	
FR	05-OCT-1999; 99US-0412863.
XX	
PA	(EPIM-) EPIMUNE INC.
XX	
PI	Settle A, Sidney J, Southwood S, Livingston BD, Cheenut R;
PI	Baker DM, Cells E, Kubo RT, Grey HM;
DR	WPI; 2001-354887/37.
XX	
XX	
PT	Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT	peptide groups, useful for vaccinating against HIV-1 -
XX	
PS	Example 3; Page 426; 448pp; English.
XX	
CC	The present invention describes a composition (I) comprising a prepared
CC	human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC	sequence selected from 51 defined amino acid sequences (AB125347 to
CC	AB25397). (I) has virucide activity and can be used in vaccines. (I)
CC	may be used for immunising subjects against HIV-1 infections. The use of
CC	group-based vaccines has several advantages over traditional vaccines,
CC	particularly when compared to the use of whole antigens in vaccine
CC	compositions. There is evidence that the immune response to whole
CC	antigens is directed largely toward variable regions of the antigen,
CC	allowing for immune escape due to mutations. The groups for inclusion in
CC	an group-based vaccine may be selected from conserved regions of viral or
CC	tumour-associated antigens, which therefore reduces the likelihood of
CC	escape mutants. Furthermore, immunosuppressive groups that may be present
CC	in whole antigens can be avoided with the use of group-based vaccines.
CC	An additional advantage of an group-based vaccine approach is the ability
CC	to combine selected groups (CTL and HTL), and further, to modify the

CC	composition of the groups, achieving, for example, enhanced	Score 52,	DB 22,	Length 10,
CC	immunogenicity. Accordingly, the immune response can be modulated, as	100.0%,		
CC	appropriate, for the target disease. Similar engineering of the response	100.0%,	Pred. No. 0.0068;	
CC	is not possible with traditional approaches. ABP1501 to ABP5412	0,	Mismatches 0,	Indels 0; Gaps 0.
CC	represent peptide sequences used in the exemplification of the present			
CC	invention.			
SO	Sequence 10 AA;			
QY	1 TTTFCASDAK 10			
DB	1 TTTFCASDAK 10			
RESULT 11				
ID	ABP25303			
XX	ABP25303 standard; Peptide; 10 AA.			
XX	ABP25303,			
XX	15-JUL-2002 (first entry)			
DE	HIV CTL epitope peptide #23.			
KM	HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vif;			
KM	vpr; vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope;			
KW	antigen; vaccine; HIV infection; immunisation; virulence.			
XX				
OS	Human immunodeficiency virus type 1.			
XX				
PN	WC000124810-A1.			
PD	12-APR-2001.			
XX				
PF	05-OCT-2000; 2000WO.0527766.			
XX				
PR	05-OCT-1999; 99US-0412863.			
XX				
PA	(EPIM-) EPIMUNE INC.			
XX				
P1	Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;			
P1	Baker DM, Celis E, Kubo RT, Grey HM;			
XX				
DR	WPI; 2001-354987/37.			
XX				
PT	Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)			
PT	peptide groups, useful for vaccinating against HIV-1 -			
XX				
PS	Example 7; Page 431; 448pp; English.			
XX				
CC	The present invention describes a composition (I) comprising a prepared			
CC	human immunodeficiency virus-1 (HIV-1) group comprising an amino acid			
CC	sequence selected from 51 defined amino acid sequences (ABU25347 to			
CC	ABR253197). (I) has virucide activity and can be used in vaccines. (I)			
CC	may be used for immunising subjects against HIV-1 infections. The use of			
CC	group-based vaccines has several advantages over traditional vaccines,			
CC	particularly when compared to the use of whole antigens in vaccine			
CC	compositions. There is evidence that the immune response to whole			
CC	antigens is directed largely toward variable regions of the antigen,			
CC	allowing for immune escape due to mutations. The groups for inclusion in			
CC	an group-based vaccine may be selected from conserved regions of viral or			
CC	tumour-associated antigens, which therefore reduces the likelihood of			
CC	escape mutants. Furthermore, immunosuppressive groups that may be present			
CC	in whole antigens can be avoided with the use of group-based vaccines.			
CC	An additional advantage of an group-based vaccine approach is the ability			
CC	to combine selected groups (CTL and HTL), and further, to modify the			
CC	composition of the groups, achieving, for example, enhanced			
CC	immunogenicity. Accordingly, the immune response can be modulated, as			
CC	appropriate, for the target disease. Similar engineering of the response			

CC is not possible with traditional approaches. ABP11501 to ABP25412  
CC represent peptide sequences used in the exemplification of the present  
XX invention.  
SQ Sequence 10 AA;  
Query Match 100.0%; Score 52; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTLCASDAK 10  
Db 1 TTLCASDAK 10  
RESULT 12  
ID ABP25363 standard; Peptide; 10 AA.  
AC ABP25363;  
XX 15-JUL-2002 (first entry)  
DE HIV-1 epitope #17.  
XX HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KM vpr; vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope;  
KM antigen; vaccine; HIV infection; immunisation; virucide.  
XX Human immunodeficiency virus type 1.  
OS WO200124810-A1.  
XX 12-APR-2001.  
PD 05-OCT-2000; 2000WO-US27766.  
XX 05-OCT-1999; 99US-0412863.  
PR (EPIM-) EPIMMUNE INC.  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX WPI; 2001-354887/37.  
DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
PT peptide groups, useful for vaccinating against HIV-1 -  
PS Claim 1; Page 434; 448pp; English.  
XX The present invention describes a composition (I) comprising a prepared  
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
CC sequence selected from 51 defined amino acid sequences (AB125347 to  
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
CC may be used for immunising subjects against HIV-1 infections. The use of  
CC group-based vaccines has several advantages over traditional vaccines,  
CC particularly when compared to the use of whole antigens in vaccine  
CC compositions. There is evidence that the immune response to whole  
CC antigens is directed largely toward variable regions of the antigen,  
CC allowing for immune escape due to mutations. The groups for inclusion in  
CC an group-based vaccine may be selected from conserved regions of viral or  
CC tumour-associated antigens, which therefore reduces the likelihood of  
CC escape mutants. Furthermore, immunosuppressive groups that may be present  
CC in whole antigens can be avoided with the use of group-based vaccines.  
CC An additional advantage of an group-based vaccine approach is the ability  
CC to combine selected groups (CTL and HTL), and further, to modify the  
CC composition of the groups, achieving, for example, enhanced  
CC immunogenicity. Accordingly, the immune response can be modulated, as  
CC appropriate, for the target disease. Similar engineering of the response  
CC is not possible with traditional approaches. ABP11501 to ABP25412  
CC represent peptide sequences used in the exemplification of the present  
CC invention.

XX SQ Sequence 10 AA;  
Query Match 100.0%; Score 52; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTLCASDAK 10  
Db 1 TTLCASDAK 10  
RESULT 13  
ID AAU96010 standard; Peptide; 10 AA.  
AC AAU96010;  
XX 02-JUL-2002 (first entry)  
DE Human immunodeficiency virus immunogenic peptide #5.  
XX Immunogenic peptide; human; major histocompatibility complex;  
KM human immunodeficiency virus; T cell activation; vaccine; HLA;  
KM viral disease; hepatitis B; Epstein-Barr; Lassa fever; papilloma;  
KM cytomegalo virus; cancer; lymphoma; prostate-specific antigen;  
KM p53; carcino-embryonal antigen; Her2/neu; autoimmune disease;  
KM human leukocyte antigen; antibody; CEA; hepatitis C; HIV.  
XX Human immunodeficiency virus type 1.  
OS WO200220053-A1.  
XX 14-MAR-2002.  
PD 01-SEP-2000; 2000WO-US24100.  
XX 01-SEP-2000; 2000WO-US24100.  
PR (EPIM-) EPIMMUNE INC.  
XX Kubo RT, Grey HM, Sette A, Celis E, Southwood S;  
PI WPI; 2002-351744/38.  
DR New immunogenic peptide, useful in vaccines against e.g. viral  
PT infection and cancer, induces a cytotoxic T cell response -  
PS Claim 1; Page 32; 39pp; English.  
XX This invention relates to a novel composition comprising an immunogenic  
CC peptide capable of specifically binding selected human major  
CC histocompatibility (MHC) class I antigens and inducing T cell  
CC activation. The peptides of the invention may also be used to induce an  
CC immune response against a desired antigen. The peptides of the invention  
CC may be used, as vaccines, to treat or prevent viral diseases  
CC (hepatitis B or C, Epstein-Barr, human immunodeficiency virus, Lassa  
CC fever, papilloma or cytomegalo viruses); cancers (e.g. of prostate,  
CC kidney or cervix, or lymphoma, where associated with expression of  
CC prostate-specific antigen, p53, carcino-embryonal antigen or Her2/neu);  
CC infection by Mycobacterium tuberculosis and autoimmune diseases. The  
CC peptides are also useful as diagnostic agents, e.g. to predict the  
CC outcome of a particular therapy and to identify subjects at risk of  
CC developing a chronic infection, also for raising specific antibodies,  
CC potentially useful as diagnostic or therapeutic agents. Nucleic acid  
CC sequences that encode the peptides can be used in DNA vaccines.  
CC The immunogenic peptides of the invention bind to HLA alleles that are  
CC widely distributed in humans. The present sequence represents an  
CC immunogenic peptide of the invention.  
SQ Sequence 10 AA;  
Query Match 100.0%; Score 52; DB 23; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
DB 1 TTLCASDAK 10

RESULT 14

AAP1791 ID AAP1791 standard; peptide; 11 AA.

XX AAP1791;

DT 25-MAY-1990 (first entry)

DE HIV-1 envelope gp120 Oligopeptide epitope.

XX AIDS; HIV-1; envelope;

XX HIV-1.

XX DE3819804-A.

XX 14-DEC-1989.

XX 10-JUN-1988; 88DE-3819804.

XX 10-JUN-1988; 88DE-3819804.

XX (ORPE) ORPEGEN MED MOLEKUL.

XX Bitt C, Heinzel W, Freibell K, Hunsmann G, Bayer H, Nick S;

XX WPI, 1989-371544/51.

XX Oligo:peptide(s) representing specific HIV-1 epitope(s) -

XX useful as vaccine components and as immunogens for raising

XX diagnostic and therapeutic antibodies.

XX Claim 1; Page 6; 29pp; German.

CC The peptide is a fragment of the HIV-1 envelope gp120, useful in raising

CC antibodies and antisera against the virus.

XX Sequence 11 AA;

QY 1 TTLCASDAK 10

DB 1 TTLCASDAK 10

RESULT 15

ID AAP12415 standard; Peptide; 11 AA.

XX AAP12415;

DT 15-JUL-2002 (first entry)

DE HIV A02 super motif env peptide #472.

XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;

XX vpr; vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope;

XX antigen; vaccine; HIV infection; immunisation; virucide.

OS Human immunodeficiency virus type 1.

XX WO200124810-A1.

XX 12 APR-2001.

XX 05-OCT-2000; 2000WC-US27766.

XX 05-OCT-1999; 99US-0412863.

XX (EPIW-) EPIMUNE INC.

XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)

XX peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 123; 448pp; English.

CC The present invention describes a composition (1) comprising a prepared

CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid

CC sequence selected from 51 defined amino acid sequences (AB125347 to

CC ABP253977). (1) has virucide activity and can be used in vaccines. (1)

CC may be used for immunising subjects against HIV-1 infections. The use of

CC group-based vaccines has several advantages over traditional vaccines,

CC particularly when compared to the use of whole antigens in vaccine

CC compositions. There is evidence that the immune response to whole

CC antigens is directed largely toward variable regions of the antigen,

CC allowing for immune escape due to mutations. The groups for inclusion in

CC an group-based vaccine may be selected from conserved regions of viral or

CC tumour-associated antigens, which therefore reduces the likelihood of

CC escape mutants. Furthermore, immunosuppressive groups that may be present

CC in whole antigens can be avoided with the use of group-based vaccines.

CC An additional advantage of an group-based vaccine approach is the ability

CC to combine selected groups (CTL and HTL), and further, to modify the

CC composition of the groups, achieving, for example, enhanced

CC immunogenicity. Accordingly, the immune response can be modulated, as

CC appropriate, for the target disease. Similar engineering of the response

CC is not possible with traditional approaches. AB11501 to ABP25412

CC represent peptide sequences used in the exemplification of the present

XX invention.

XX Sequence 11 AA;

QY 1 TTLCASDAK 10

DB 1 TTLCASDAK 10

Search completed: April 14, 2003, 06:38:01  
Job time : 39 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 06:39:07 ; Search time 343 Seconds  
(without alignments)  
1.762 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTLEFASDAK 10

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodaca/1/pubppaa/US08\_NEW\_PUB\_PEP.\*
- 2: /cgn2\_6/ptodaca/1/pubppaa/US05\_NEW\_PUB\_PEP.\*
- 3: /cgn2\_6/ptodaca/1/pubppaa/US06\_NEW\_PUB\_PEP.\*
- 4: /cgn2\_6/ptodaca/1/pubppaa/US07\_NEW\_PUB\_PEP.\*
- 5: /cgn2\_6/ptodaca/1/pubppaa/US08\_NEW\_PUB\_PEP.\*
- 6: /cgn2\_6/ptodaca/1/pubppaa/US09\_NEW\_PUB\_PEP.\*
- 7: /cgn2\_6/ptodaca/1/pubppaa/US10\_NEW\_PUB\_PEP.\*
- 8: /cgn2\_6/ptodaca/1/pubppaa/US05\_PUBCOMB\_PEP.\*
- 9: /cgn2\_6/ptodaca/1/pubppaa/US06\_PUBCOMB\_PEP.\*
- 10: /cgn2\_6/ptodaca/1/pubppaa/US07\_PUBCOMB\_PEP.\*
- 11: /cgn2\_6/ptodaca/1/pubppaa/US08\_PUBCOMB\_PEP.\*
- 12: /cgn2\_6/ptodaca/1/pubppaa/US09\_PUBCOMB\_PEP.\*
- 13: /cgn2\_6/ptodaca/1/pubppaa/US10\_PUBCOMB\_PEP.\*
- 14: /cgn2\_6/ptodaca/1/pubppaa/US60\_PUBCOMB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	1 US-08-821-739A-89	Sequence 89, Appl
2	52	100.0	10	10 US-09-894-018-164	Sequence 164, App
3	52	100.0	12	10 US-09-911-838-2	Sequence 2, Appl
4	52	100.0	100	9 US-10-032-162-2	Sequence 85, Appl
5	52	100.0	211	10 US-09-894-018-85	Sequence 81, Appl
6	52	100.0	277	10 US-09-894-018-81	Sequence 83, Appl
7	52	100.0	280	10 US-09-894-018-83	Sequence 16, Appl
8	52	100.0	502	10 US-09-796-202-16	Sequence 17, Appl
9	52	100.0	503	10 US-09-759-841-4	Sequence 48, Appl
10	52	100.0	511	10 US-09-796-202-17	Sequence 15, Appl
11	52	100.0	516	9 US-10-026-741-48	Sequence 87, Appl
12	52	100.0	579	9 US-10-032-162-15	Sequence 4, Appl
13	52	100.0	585	10 US-09-894-018-87	Sequence 17, Appl
14	52	100.0	619	10 US-09-891-609-4	Sequence 13, Appl
15	52	100.0	625	9 US-10-032-162-17	Sequence 2, Appl
16	52	100.0	643	9 US-10-032-162-13	Sequence 19, Appl
17	52	100.0	646	10 US-09-891-609-2	Sequence 2, Appl
18	52	100.0	844	10 US-09-991-258-19	Sequence 2, Appl
19	52	100.0	847	10 US-09-476-242-2	Sequence 2, Appl

20	52	100.0	853	9 US-10-003-035-33	Sequence 33, Appl
21	52	100.0	856	10 US-09-476-242-1	Sequence 1, Appl
22	52	100.0	861	9 US-10-026-741-103	Sequence 103, App
23	52	100.0	868	9 US-09-938-406-1	Sequence 1, Appl
24	52	100.0	1101	9 US-10-003-035-53	Sequence 53, Appl
25	52	100.0	1185	9 US-09-819-401-18	Sequence 18, Appl
26	52	100.0	1186	9 US-10-003-035-55	Sequence 55, Appl
27	47	90.4	9	1 US-08-821-739A-67	Sequence 67, Appl
28	47	90.4	506	9 US-09-934-060A-24	Sequence 24, Appl
29	47	90.4	506	9 US-09-934-060A-10	Sequence 30, Appl
30	47	90.4	556	9 US-09-934-060A-6	Sequence 6, Appl
31	47	90.4	720	9 US-09-934-060A-2	Sequence 2, Appl
32	47	90.4	720	9 US-09-934-060A-4	Sequence 4, Appl
33	42	80.8	21	9 US-10-062-710-155	Sequence 155, App
34	42	80.8	526	9 US-10-026-741-46	Sequence 46, Appl
35	42	80.8	877	9 US-10-026-741-102	Sequence 102, App
36	38	73.1	7	10 US-09-911-838-53	Sequence 53, Appl
37	37	71.2	7	10 US-09-911-838-57	Sequence 57, Appl
38	37	71.2	7	10 US-09-911-838-51	Sequence 51, Appl
39	37	71.2	7	10 US-09-911-838-55	Sequence 55, Appl
40	37	71.2	873	1 US-08-911-824-61	Sequence 61, Appl
41	36	69.2	514	10 US-09-886-468-23	Sequence 23, Appl
42	33	63.5	6	10 US-09-911-838-50	Sequence 50, Appl
43	33	63.5	6	10 US-09-911-838-54	Sequence 54, Appl
44	33	63.5	6	10 US-09-911-838-56	Sequence 56, Appl
45	33	63.5	7	10 US-09-911-838-49	Sequence 49, Appl

#### ALIGNMENTS

RESULT 1  
US-08-821-739A-89  
Sequence 89, Application US/08821739A  
Patent No. US20020168374A1  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Bjarne  
TITLE OF INVENTION: HLA Binding Peptides and Their Uses  
FILE REFERENCE: 2060.005000A  
CURRENT APPLICATION NUMBER: US/08/821,739A  
CURRENT FILING DATE: 1999-03-20  
PRIOR APPLICATION NUMBER: 60/013,833  
PRIOR FILING DATE: 1996-03-21  
PRIOR APPLICATION NUMBER: 08/589,107  
PRIOR FILING DATE: 1996-07-12  
PRIOR APPLICATION NUMBER: 08/451,913  
PRIOR FILING DATE: 1995-05-26  
PRIOR APPLICATION NUMBER: 08/347,610  
PRIOR FILING DATE: 1994-12-01  
PRIOR APPLICATION NUMBER: 08/186,266  
PRIOR FILING DATE: 1994-01-25  
PRIOR APPLICATION NUMBER: 08/159,339  
PRIOR FILING DATE: 1993-11-29  
PRIOR APPLICATION NUMBER: 08/103,396  
PRIOR FILING DATE: 1993-08-06  
PRIOR APPLICATION NUMBER: 08/027,746  
PRIOR FILING DATE: 1993-03-05  
PRIOR APPLICATION NUMBER: 07/926,666  
PRIOR FILING DATE: 1992-08-07  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 89  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-821-739A-89  
Query Match 100.0%; Score 52; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 1 TTLCASDAK 10

RESULT 2  
US-09-894-018-164  
; Sequence 164, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian  
; APPLICANT: Baker, Denisw  
; APPLICANT: Newman, Mark  
; APPLICANT: Brown, David  
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY  
; FILE REFERENCE: 39963-20033.00  
; CURRENT APPLICATION NUMBER: US/09/894,018  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: PCT/US00/35568  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: US 60/173,390  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 60/284,221  
; PRIOR FILING DATE: 2001-04-16  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 164  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Transgenic mouse  
; US-09-894-018-164

Query Match 100.0%; Score 52; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 1 TTLCASDAK 10

RESULT 3  
US-09-911-838-2  
; Sequence 2, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; TITLE OF INVENTION: SYNROME  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT APPLICATION NUMBER: US/09/911,838  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-911-838-2

Query Match 100.0%; Score 52; DB 10; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 1 TTLCASDAK 10

RESULT 4  
US-10-032-162-2  
; Sequence 2, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331a2  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-2

Query Match 100.0%; Score 52; DB 9; Length 100;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
| | | | |  
Db 20 TTLCASDAK 29

RESULT 5  
US-09-894-018-85  
; Sequence 85, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian  
; APPLICANT: Baker, Denisw  
; APPLICANT: Newman, Mark  
; APPLICANT: Brown, David  
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY  
; FILE REFERENCE: 39963-20033.00  
; CURRENT APPLICATION NUMBER: US/09/894,018  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: PCT/US00/35568  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: US 60/173,390  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 60/284,221  
; PRIOR FILING DATE: 2001-04-16  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 85  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HIV-FT  
US-09-894-018-85

Query Match 100.0%; Score 52; DB 10; Length 211;

Best Local Similarity 100.0%; Pred. No. 0.036;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 88 TTLCASDAK 97

## RESULT 6

US-09-894-018-81  
; Sequence 81, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian  
; APPLICANT: Baker, Denisw  
; APPLICANT: Newman, Mark  
; APPLICANT: Brown, David  
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
; FILE OF INVENTION: MINIGENES AND PEPTIDES THEREBY  
; FILE REFERENCE: 39963-20033.00  
; CURRENT APPLICATION NUMBER: US/09/894,018  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: PCT/US00/35568  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: US 60/173,390  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 60/284,221  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HIV-1090  
US-09-894-018-81

Query Match 100.0%; Score 52; DB 10; Length 277;

Best Local Similarity 100.0%; Pred. No. 0.047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 141 TTLCASDAK 150

## RESULT 7

US-09-894-018-83  
; Sequence 83, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian  
; APPLICANT: Baker, Denisw  
; APPLICANT: Newman, Mark  
; APPLICANT: Brown, David  
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
; FILE OF INVENTION: MINIGENES AND PEPTIDES THEREBY  
; FILE REFERENCE: 39963-20033.00  
; CURRENT APPLICATION NUMBER: US/09/894,018  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: PCT/US00/35568  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: US 60/173,390  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 60/284,221  
; PRIOR FILING DATE: 2001-04-16

NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 83  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HIV-CPT  
US-09-894-018-83

## Query Match

Best Local Similarity 100.0%; Pred. No. 0.048;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 238 TTLCASDAK 247

## RESULT 8

US-09-796-202-16  
; Sequence 16, Application US/09796202  
; Patent No. US20020068813A1  
; GENERAL INFORMATION:  
; APPLICANT: Dragic, Tatjana  
; APPLICANT: Olson, William  
; TITLE OF INVENTION: SUPRATED CCR5 PEPTIDES FOR HIV-1 INFECTION  
; FILE REFERENCE: 2048/61010/JPM/SHS  
; CURRENT APPLICATION NUMBER: US/09/796,202  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: human immunodeficiency virus  
US-09-796-202-16

Query Match 100.0%; Score 52; DB 10; Length 502;

Best Local Similarity 100.0%; Pred. No. 0.084;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTLCASDAK 10  
Db 49 TTLCASDAK 58

## RESULT 9

US-09-759-841-4  
; Sequence 4, Application US/09759841  
; Patent No. US20010039026A1  
; GENERAL INFORMATION:  
; APPLICANT: Rickett, Graham A  
; APPLICANT: Dobbs, Susan  
; APPLICANT: Petros, Manousos  
; TITLE OF INVENTION: Assay Method  
; FILE REFERENCE: PC10348APME  
; CURRENT APPLICATION NUMBER: US/09/759,841  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: GB 0000661.9  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: GB 0000663.5  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: GB 0000659.3  
; PRIOR FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-759-841-4

Query Match 100.0%; Score 52; DB 10; Length 503;  
Best Local Similarity 100.0%; Pred. No. 0.084;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 43 TTLCASDAK 52

## RESULT 10

US-09-796-202-17  
; Sequence 17, Application US/09796202  
; Patent No. US20020068813A1  
; GENERAL INFORMATION:  
; APPLICANT: Dragic, Tatjana  
; APPLICANT: Olson, William  
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION  
; FILE REFERENCE: 2048/61010/JPM/SHS  
; CURRENT APPLICATION NUMBER: US/09/796,202  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: human immunodeficiency virus  
US-09-796-202-17

Query Match 100.0%; Score 52; DB 10; Length 511;  
Best Local Similarity 100.0%; Pred. No. 0.085;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 50 TTLCASDAK 59

## RESULT 11

US-10-026-741-48  
; Sequence 48, Application US/10026741  
; Publication No. US20030049604A1  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; CLAVEL, FRANCOISE  
; BORMAN, ANDRE  
; OUILLENT, CAROLINE  
; GUEYARD, DENISE  
; MONTAGNIER, LUC  
; DONJON DE SAINT-MARTIN, JACQUELINE  
; COHEN, JAOUES  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; SUBTYPE) ANTIGENS  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N W  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/026,741  
; FILING DATE: 27-Dec-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/817,441  
; FILING DATE: 31-AUG-1998  
; APPLICATION NUMBER: PCT/FR 95/01391

FILING DATE: 20-OCT-1995  
APPLICATION NUMBER: FR 9412554  
FILING DATE: 20-OCT-1994  
APPLICATION NUMBER: FR 9502526  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 516 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:

Query Match 100.0%; Score 52; DB 9; Length 516;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 50 TTLCASDAK 59

## RESULT 12

US-10-032-162-15  
; Sequence 15, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331a2  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-15

Query Match 100.0%; Score 52; DB 9; Length 579;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 20 TTLCASDAK 29

## RESULT 13

US-09-894-018-87  
; Sequence 87, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian

APPLICANT: Baker, Denisw  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David  
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894,018  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/284,221  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 87  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV-TC  
US-09-894-018-87

Query Match 100.0%; Score 52; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 0.097;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 121 TTLCASDAK 130

RESULT 14  
US-09-891-609-4  
Sequence 4, Application US/09891609  
Patent No. US20020127238A1  
GENERAL INFORMATION:  
APPLICANT: Stamatos, Leonidas  
TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
FILE REFERENCE: 2570-1-002N  
CURRENT APPLICATION NUMBER: US/09/891,609  
CURRENT FILING DATE: 2001-06-26  
PRIOR APPLICATION NUMBER: 60/214,608  
PRIOR FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 619  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-4

Query Match 100.0%; Score 52; DB 10; Length 619;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 22 TTLCASDAK 31

RESULT 15  
US-10-032-162-17  
Sequence 17, Application US/10032162  
Publication No. US20030052839A1  
GENERAL INFORMATION:  
APPLICANT: BINEY, JAMES M  
APPLICANT: SCHUELEKE, NORBERT  
APPLICANT: OLSON, WILLIAM C  
APPLICANT: PAUL, MADDON J  
APPLICANT: JOHN, MOORE P  
TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF

FILE REFERENCE: 2048/59331a2  
CURRENT APPLICATION NUMBER: US/10/032,162  
CURRENT FILING DATE: 2002-10-24  
PRIOR APPLICATION NUMBER: 09/602,864  
PRIOR FILING DATE: 2000-06-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 625  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (151)..(151)  
OTHER INFORMATION: X=UNKNOWN AMINO ACID  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (205)..(205)  
OTHER INFORMATION: X=UNKNOWN AMINO ACID  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (213)..(213)  
OTHER INFORMATION: X=UNKNOWN AMINO ACID  
US-10-032-162-17

Query Match 100.0%; Score 52; DB 9; Length 625;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 20 TTLCASDAK 29

Search completed: April 14, 2003, 07:02:53  
Job time : 344 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

## OM protein - protein search, using sw model

Run on: April 14, 2003, 06:37:13 ; Search time 14 Seconds

(without alignments)  
21.016 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTLCASDAK 10

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilist1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	20	1 US-08-218-025A-34	Sequence 34, Appl
2	52	100.0	20	1 US-08-306-116A-3	Sequence 3, Appl
3	52	100.0	24	2 US-08-493-235-3	Sequence 3, Appl
4	52	100.0	32	5 PCT-US91-02166-9	Sequence 9, Appl
5	52	100.0	455	3 US-08-889-841B-46	Sequence 46, Appl
6	52	100.0	474	3 US-08-889-841B-36	Sequence 36, Appl
7	52	100.0	474	3 US-08-889-841B-39	Sequence 39, Appl
8	52	100.0	479	2 US-08-037-816A-20	Sequence 20, Appl
9	52	100.0	479	2 US-08-037-816A-24	Sequence 24, Appl
10	52	100.0	479	2 US-08-530-146-20	Sequence 20, Appl
11	52	100.0	479	2 US-08-530-146-24	Sequence 24, Appl
12	52	100.0	479	5 PCT-US91-02166-10	Sequence 10, Appl
13	52	100.0	479	5 PCT-US91-02250-1	Sequence 1, Appl
14	52	100.0	483	3 US-08-889-841B-31	Sequence 31, Appl
15	52	100.0	484	2 US-08-037-816A-22	Sequence 22, Appl
16	52	100.0	484	2 US-08-530-146-22	Sequence 22, Appl
17	52	100.0	486	3 US-08-889-841B-8	Sequence 8, Appl
18	52	100.0	487	3 US-08-889-841B-33	Sequence 33, Appl
19	52	100.0	491	2 US-08-037-816A-18	Sequence 18, Appl
20	52	100.0	491	2 US-08-530-146-18	Sequence 18, Appl
21	52	100.0	491	3 US-08-889-841B-10	Sequence 10, Appl
22	52	100.0	495	3 US-08-889-841B-25	Sequence 25, Appl
23	52	100.0	495	3 US-08-889-841B-28	Sequence 28, Appl
24	52	100.0	496	3 US-08-889-841B-12	Sequence 12, Appl
25	52	100.0	496	3 US-08-889-841B-16	Sequence 16, Appl
26	52	100.0	498	2 US-07-916-098A-4	Sequence 4, Appl
27	52	100.0	498	3 US-08-889-841B-2	Sequence 2, Appl

28	52	100.0	498	3 US-08-889-841B-5	Sequence 5, Appl
29	52	100.0	501	2 US-08-448-603A-2	Sequence 2, Appl
30	52	100.0	501	3 US-09-134-078-2	Sequence 2, Appl
31	52	100.0	501	4 US-09-492-739-2	Sequence 2, Appl
32	52	100.0	507	2 US-08-037-816A-16	Sequence 16, Appl
33	52	100.0	507	2 US-08-037-816A-28	Sequence 28, Appl
34	52	100.0	507	2 US-08-530-146-16	Sequence 16, Appl
35	52	100.0	507	2 US-08-530-146-28	Sequence 28, Appl
36	52	100.0	508	4 US-08-472-240A-16	Sequence 16, Appl
37	52	100.0	511	2 US-08-448-603A-1	Sequence 1, Appl
38	52	100.0	511	3 US-09-134-078-1	Sequence 1, Appl
39	52	100.0	511	3 US-08-889-841B-41	Sequence 41, Appl
40	52	100.0	511	4 US-09-492-739-1	Sequence 1, Appl
41	52	100.0	513	4 US-08-472-240A-14	Sequence 14, Appl
42	52	100.0	516	4 US-08-817-441-48	Sequence 48, Appl
43	52	100.0	520	2 US-08-037-816A-14	Sequence 14, Appl
44	52	100.0	520	2 US-08-037-816A-26	Sequence 26, Appl
45	52	100.0	520	2 US-08-530-146-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-08-218-025A-34

; Sequence 34, Application US/08218025A

; Patent No. 5556744

; GENERAL INFORMATION:

; APPLICANT: Weiner, David B.

; APPLICANT: Ugen, Kenneth E.

; TITLE OF INVENTION: Williams, William V.

; TITLE OF INVENTION: Methods and Compositions for Diagnosing

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: P.O. Box 457, 321 No. 5556744istown Road

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/218, 025A

; FILING DATE: 24-MAR-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/891,451

; FILING DATE: 29-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: WST33A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 540-9206

; TELEFAX: (215) 540-5818

; INFORMATION FOR SEQ ID NO: 34:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; US-08-218-025A-34

Query Match 100.0%; Score 52; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10

Db 9 TTLCASDAK 18

RESULT 2

US-08-306-116A-3

Sequence 3, Application US/08306116A

Patent No. 5691135

GENERAL INFORMATION:

APPLICANT: Braun, Jonathan

APPLICANT: Goodlick, Lee A.

TITLE OF INVENTION: IMMUNOGLOBULIN SUPRANTIGEN BINDING TO GP120

TITLE OF INVENTION: FROM HIV

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive 16th floor

CITY: Newport Beach

STATE: CA

COUNTRY: U S A

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/306,116A

FILING DATE: 14-SEP-1994

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/259,669

FILING DATE: 14-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: DCLA004,001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: 714-760-0404

TELEFAX: 714-760-9502

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

US-08-306-116A-3

Query Match

Best Local Similarity 100.0%; Score 52; DB 1; Length 20;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 TTLCASDAK 18

RESULT 3

US-08-493-235-3

Sequence 3, Application US/08493235

Patent No. 5840313

GENERAL INFORMATION:

APPLICANT: Vahne, Anders

APPLICANT: Svennerholm, Bo

APPLICANT: Rybo, Lars

APPLICANT: Jeansson, Stig

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

APPLICANT: Horal, Peter

TITLE OF INVENTION: PEPTIDES FOR USE IN VACCINATION AND INDUCTION OF NEUTRALIZING ANTIBODIES AGAINST HUMAN

TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

STREET: 620 NEWPORT CENTER DRIVE 16TH FLOOR

CITY: NEWPORT BEACH

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/493,235

FILING DATE: 20 June 1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Kaiser, Annemarie

REGISTRATION NUMBER: 37,649

REFERENCE/DOCKET NUMBER: METRICS.12CPCI

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-493-235-3

Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 24;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 TTLCASDAK 19

RESULT 4

PCT-US91-02166-9

Sequence 9, Application PC/TUS9102166

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

TITLE OF INVENTION: HIV Envelope Polypeptides

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02166

FILING DATE: 19910401

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S.S.N. 07/504,772

FILING DATE: 03-APRIL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Carolyn R.  
REGISTRATION NUMBER: 32,324  
REFERENCE/DOCKET NUMBER: 639  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/266-2614  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
PCT-US91-02166-9

Query Match 100.0%; Score 52; DB 5; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 2 TTLCASDAK 11

RESULT 5  
US-08-889-841B-46  
Sequence 46, Application US/08889841B  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
CURRENT FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: US 60/676,737  
PRIOR FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 46  
LENGTH: 455  
TYPE: PRT  
ORGANISM: HIV  
US-08-889-841B-46

Query Match 100.0%; Score 52; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 9 TTLCASDAK 18

RESULT 6  
US-08-889-841B-36  
Sequence 36, Application US/08889841B  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
CURRENT FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: US 60/676,737  
PRIOR FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 36  
LENGTH: 474  
TYPE: PRT  
ORGANISM: HIV  
US-08-889-841B-36

Query Match 100.0%; Score 52; DB 3; Length 474;

Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 11 TTLCASDAK 20

RESULT 7  
US-08-889-841B-39  
Sequence 39, Application US/08889841B  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
CURRENT FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: US 60/676,737  
PRIOR FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 39  
LENGTH: 474  
TYPE: PRT  
ORGANISM: HIV  
US-08-889-841B-39

Query Match 100.0%; Score 52; DB 3; Length 474;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 11 TTLCASDAK 20

RESULT 8  
US-08-037-816A-20  
Sequence 20, Application US/08037816A  
Patent No. 5669624  
GENERAL INFORMATION:  
APPLICANT: Hasel, Karl W.  
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED  
THEREOF, AND THERAPEUTIC AND PROPHYLACTIC USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/037,816A  
FILING DATE: 26-MAR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41190/JPW/AM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOPUI  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-037-816A-20

Query Match 100.0%; Score 52; DB 2; Length 479;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 55 TTLCASDAK 64

RESULT 9  
US-08-037-816A-24  
Sequence 24, Application US/08037816A  
Patent No. 5869624  
GENERAL INFORMATION:  
APPLICANT: Hasel, Karl W.  
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED  
TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/037,816A  
FILING DATE: 26-MAR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOPUI  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 479 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-037-816A-24  
Query Match 100.0%; Score 52; DB 2; Length 479;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 55 TTLCASDAK 64

RESULT 10  
US-08-530-146-20  
Sequence 20, Application US/08530146  
Patent No. 5886163  
GENERAL INFORMATION:  
APPLICANT: Hasel, Karl W.

APPLICANT: Maddon, Paul J.  
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED  
TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/530,146  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/037,816  
FILING DATE: 26-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOPUI  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 479 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-530-146-20

Query Match 100.0%; Score 52; DB 2; Length 479;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTLCASDAK 10  
Db 55 TTLCASDAK 64

RESULT 11  
US-08-530-146-24  
Sequence 24, Application US/08530146  
Patent No. 5886163  
GENERAL INFORMATION:  
APPLICANT: Hasel, Karl W.  
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED  
TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/530,146  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/037,816  
FILING DATE: 26-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41190/UPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOPUI  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 479 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-530-146-24

Query Match 100.0%; Score 52; DB 2; Length 479;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLCASDAK 10  
Db 55 TTLCASDAK 64

RESULT 12  
PCT-US91-02166-10  
Sequence 10, Application PC/TUS9102166  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: HIV Envelope Polypeptides  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02166  
FILING DATE: 19910401  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S.S.N. 07/504,772  
FILING DATE: 03-APRIL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Carolyn R.  
REGISTRATION NUMBER: 32,324  
REFERENCE/DOCKET NUMBER: 639  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/266-2614  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 479 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
PCT-US91-02166-10

Query Match 100.0%; Score 52; DB 5; Length 479;

Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLCASDAK 10  
Db 20 TTLCASDAK 29

RESULT 13  
PCT-US91-02250-1  
Sequence 1, Application PC/TUS9102250  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: Methods and Compositions for Vaccination Against HIV  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02250  
FILING DATE: 19910401  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S.S.N. 07/504,785  
FILING DATE: 03-APRIL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Carolyn R.  
REGISTRATION NUMBER: 32,324  
REFERENCE/DOCKET NUMBER: 633  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/266-2614  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 479 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
PCT-US91-02250-1

Query Match 100.0%; Score 52; DB 5; Length 479;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLCASDAK 10  
Db 20 TTLCASDAK 29

RESULT 14  
US-08-889-841B-31  
Sequence 31, Application US/08889841B  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
PRIOR FILING DATE: 1997-07-08  
CURRENT APPLICATION NUMBER: US 60/676,737  
PRIOR FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 31  
LENGTH: 483

TYPE: PRT  
ORGANISM: HIV  
US-08-889-841B-31

Query Match 100.0%; Score 52; DB 3; Length 483;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 9 TTLCASDAK 18

RESULT 15  
US-08-037-816A-22

Sequence 22, Application US/08037816A  
Patent No. 5869624  
GENERAL INFORMATION:  
APPLICANT: Hasel, Karl W.  
APPLICANT: Maddon, Paul J.  
TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED  
TITLE OF INVENTION: THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/037, 816A  
FILING DATE: 26-MAR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41190/JPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-037-816A-22

Query Match 100.0%; Score 52, DB 2, Length 484,  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCASDAK 10  
Db 55 TTLCASDAK 64

Search completed: April 14, 2003, 06:39:21  
Job time : 16 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 06:57:09 ; Search time 42 seconds  
(without alignments)  
22.889 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTLFCASDAK 10

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613442 residues

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	40.4	10	2 PH0927	T-cell receptor be
2	20	38.5	6	2 137263	Y protein - human
3	19	36.5	7	2 PH1408	Ig heavy chain V r
4	19	36.5	8	2 PH1407	Ig heavy chain V r
5	19	36.5	10	2 139702	nopaline synthase
6	19	36.5	10	2 PH0933	T-cell receptor be
7	17	32.7	7	2 PH1602	Ig H chain V-D-J r
8	17	32.7	7	2 PH0932	T-cell receptor be
9	17	32.7	8	2 PH0934	T-cell receptor be
10	17	32.7	9	2 PH0943	T-cell receptor be
11	17	32.7	9	2 PH0935	T-cell receptor be
12	17	32.7	9	2 PH0937	T-cell receptor be
13	17	32.7	9	2 PH0930	T-cell receptor be
14	17	32.7	9	2 PH0917	T-cell receptor be
15	17	32.7	9	2 PH0918	T-cell receptor be
16	17	32.7	9	2 PH0921	T-cell receptor be
17	17	32.7	10	2 D61440	polygalacturonase
18	17	32.7	10	2 PT0215	T-cell receptor be
19	17	32.7	10	2 PH0900	T-cell receptor be
20	17	32.7	10	2 PH0944	T-cell receptor be
21	17	32.7	10	2 PH0894	T-cell receptor be
22	17	32.7	10	2 PH0946	T-cell receptor be
23	17	32.7	10	2 PH0925	T-cell receptor be
24	17	32.7	10	2 PH0926	T-cell receptor be
25	17	32.7	10	2 PH0916	T-cell receptor be
26	17	32.7	10	2 PH0923	T-cell receptor be
27	17	32.7	10	2 PH0895	T-cell receptor be
28	17	32.7	10	2 PH0948	T-cell receptor be
29	16	30.8	4	2 S43959	Ig mu chain V regi

30	16	30.8	6	2 P41946	T-cell receptor ga
31	16	30.8	9	2 I53874	seminal vesicle pr
32	16	30.8	10	2 E41946	T-cell receptor ga
33	16	30.8	10	2 C41946	T-cell receptor ga
34	15	28.8	5	2 PT0596	T-cell receptor be
35	15	28.8	6	2 PT0616	T-cell receptor be
36	15	28.8	6	2 PT0650	T-cell receptor be
37	15	28.8	6	2 PT0650	T-cell receptor be
38	15	28.8	6	2 PT0693	T-cell receptor be
39	15	28.8	7	2 PT0524	T-cell receptor be
40	15	28.8	7	2 PT0521	T-cell receptor be
41	15	28.8	7	2 PT0683	T-cell receptor be
42	15	28.8	8	2 PT0627	T-cell receptor be
43	15	28.8	8	2 PT0509	T-cell receptor be
44	15	28.8	8	2 PT0547	T-cell receptor be
45	15	28.8	8	2 PT0716	T-cell receptor be

#### ALIGNMENTS

RESULT 1  
PH0927  
T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0927  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
U. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
A:Reference number: PH0891, MUID:92078857, PMID:1836012  
A:Accession: PH0927  
A:Molecule type: mRNA  
A:Residues: 1-10 <GOL>  
A:Experimental source: concanavalin A-activated lymphoblast  
C:Keywords: T-cell receptor

Query Match 40.4%; Score 21; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 8.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CASDA 9  
Db 1 CASSA 5

RESULT 2  
Y protein - human (fragment)  
137263  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: 137263  
R:Maeder, G.; Habener, J.F.  
Endocrinology 131, 2010-2015, 1992  
A:Title: Novel testis germ cell-specific transcript of the CREB gene contains an alterna  
A:Reference number: 137263; MUID:93010691, PMID:1396344  
A:Accession: 137263  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: EMBL:X68994; NID:9396171; PIDN:CAA48780.1; PID:9579816  
C:Genetics:  
A:Gene: CREB

Query Match 38.5%; Score 20; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLFC 5  
Db 1 SLFC 4



A:Experimental source: complete Freund's adjuvant-immunized lymph node  
C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 9

PH0934

T-cell receptor beta chain V-D-J region (clone 5) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0934

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0934

A:Molecule type: mRNA

A:Residues: 1-8 <GOL>

A:Experimental source: complete Freund's adjuvant-immunized lymph node

A>Note: the authors translated the codon CAG for residue 7 as Glu

C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 10

PH0943

T-cell receptor beta chain V-D-J region (clone 14) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0943

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0943

A:Molecule type: mRNA

A:Residues: 1-9 <GOL>

A:Experimental source: complete Freund's adjuvant-immunized lymph node

A>Note: the authors translated the codon GAC for residue 8 as Glu

C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 11

PH0935

T-cell receptor beta chain V-D-J region (clone 6) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0935

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0935

A:Molecule type: mRNA

A:Residues: 1-9 <GOL>

A:Experimental source: complete Freund's adjuvant-immunized lymph node

C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 12

PH0937

T-cell receptor beta chain V-D-J region (clone 8) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0937

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0937

A:Molecule type: mRNA

A:Residues: 1-9 <GOL>

A:Experimental source: complete Freund's adjuvant-immunized lymph node

C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 13

PH0902

T-cell receptor beta chain V-D-J region (hybridoma S23F4F4) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0902

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0902

A:Molecule type: mRNA

A:Residues: 1-9 <GOL>

A:Experimental source: myelin basic protein-immunized T-cell

C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAS 7  
|||  
DB 1 CAS 3

## RESULT 14

PH0917

T-cell receptor beta chain V-D-J region (isolate 3) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997

C:Accession: PH0917

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
A:Reference number: PH0891; MUID:92078857; PMID:1836012  
A:Accession: PH0918  
A:Molecule type: mRNA  
A:Residues: 1-9 <GOL>  
A:Experimental source: concanavalin A-activated lymphoblast  
C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

OY 5 CAS 7  
|||  
Db 1 CAS 3

## RESULT 15

PH0918  
T-cell receptor beta chain V-D-J region (isolate 4) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0918  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergy  
A:Reference number: PH0891; MUID:92078857; PMID:1836012  
A:Accession: PH0918  
A:Molecule type: mRNA  
A:Residues: 1-9 <GOL>  
A:Experimental source: concanavalin A-activated lymphoblast  
C:Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

OY 5 CAS 7  
|||  
Db 1 CAS 3

Search completed: April 14, 2003, 07:05:55  
Job time : 43 secs

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 06:39:26 ; Search time 11 seconds

(without alignments) 37.706 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTTFCASDAK 10

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 346

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	15	28.8	9 1 CCAP_CARMA	P38556 carcinius ma
2	15	28.8	9 1 SAP_STOVA	P24547 stomopneute
3	14	26.9	9 1 UPB3_HUMAN	P30089 homo sapien
4	14	26.9	10 1 FIBB_CERSI	P14537 ceratocheri
5	14	26.9	10 1 URA6_HUMAN	P32080 homo sapien
6	14	26.9	10 1 UXA2_CHLTR	P38003 chlamydia t
7	13	25.0	10 1 UPB9_HUMAN	P30095 homo sapien
8	12	23.1	7 1 UC24_MAIZE	P80630 zea mays (m
9	12	23.1	8 1 ACT_CARMA	P80709 carcinius ma
10	12	23.1	8 1 ORMY_ORCLI	P82455 orconectes
11	12	23.1	10 1 COXA_ONCMY	P80328 oncorhynch
12	12	23.1	10 1 COXO_SHEEP	P80337 ovis aries
13	11	21.2	7 1 CCPI_ENTFA	P20104 enterococcu
14	11	21.2	8 1 FARB_CALVO	P41863 calliphora
15	11	21.2	9 1 DSIP_RABIT	P01158 oryctolagus
16	11	21.2	9 1 FIBB_MACFU	P19345 macaca fusc
17	11	21.2	9 1 MOSH_CLYJA	P19853 clypeaster
18	11	21.2	9 1 MOSH_CLYJA	P19852 clypeaster
19	11	21.2	10 1 MALE_KLEBN	P05564 klebsiella
20	11	21.2	10 1 MOSO_CLYJA	P19862 clypeaster
21	11	21.2	10 1 TKNB_RANCA	P26869 rana catesb
22	11	21.2	10 1 TPIS_NICPL	P19118 nicotiana p
23	11	21.2	5 1 B10A_CITFR	P13071 citrobacter
24	10	19.2	7 1 BRHP_CONIM	P58803 conus imper
25	10	19.2	8 1 ALIB_CARMA	P13368 carcinius ma
26	10	19.2	8 1 CAD1_ENTFA	P13368 enterococcu
27	10	19.2	8 1 COM2_CONPU	P58785 conus purpu
28	10	19.2	8 1 FARI_PANRE	P41672 panagrellus
29	10	19.2	8 1 FAR3_HOMAM	P41486 homarus ame
30	10	19.2	8 1 LCK4_LEUMA	P21143 leucophaea
31	10	19.2	8 1 NPB_BOVIN	P15507 bos taurus
32	10	19.2	9 1 COM_CONVE	P83047 conus ventr
33	10	19.2	9 1 FIBB_ERYPFA	P19346 erythrocebu

34	10	19.2	9 1 FIBB_PAPNA	P19344 papio anubi
35	10	19.2	9 1 FIBB_PAPHA	P19343 papio hamad
36	10	19.2	9 1 FIBB_THBGE	P19342 theopithec
37	10	19.2	9 1 HUTU_KLEAE	P12381 klebsiella
38	10	19.2	9 1 IPYR_RHOVI	P82992 rhodopsu
39	10	19.2	9 1 NEUV_CAVPO	P34966 cavia porce
40	10	19.2	9 1 TAL1_PICUA	P17440 picchia jadi
41	10	19.2	9 1 TAL3_PICUA	P17441 picchia jadi
42	10	19.2	9 1 TKCI_CALVO	P41517 calliphora
43	10	19.2	9 1 ULAE_HUMAN	P31931 homo sapien
44	10	19.2	9 1 ULAE_HUMAN	P31934 homo sapien
45	10	19.2	10 1 AMPN_HELAM	P81731 helicoverpa

## ALIGNMENTS

## RESULT 1

CCAP\_CARMA STANDARD; PRT; 9 AA.

AC P38556; 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Cardioactive peptide (CCAP).  
OS Carcinus maenas (Common shore crab) (Green crab),  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm),  
OS Tenebrio molitor (Yellow mealworm), and  
OS Spodoptera eridania (Southern armyworm).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
OC Brachyura; Eubrachyura; Portunidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759, 7130, 7067, 37547,  
[1]

## SEQUENCE.

RC SPECIES=C.maenas; TISSUE=Pericardial organs;  
RA Stangier J., Hilbich C., Beyreuther K., Keller R.;  
RT "Unusual cardioactive peptide (CCAP) from pericardial organs of the  
RT shore crab Carcinus maenas";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:575-579(1987).  
[2]

RC SPECIES=M.sexta;  
RX MEDLINE=93050243; PubMed=1426284;  
RA Cheung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;  
RT "Primary structure of a cardioactive neuropeptide from the tobacco  
RT hawkmoth, Manduca sexta";  
RL FEBS Lett. 313:165-168(1992).  
[3]

## SEQUENCE.

RC SPECIES=T.molitor, and S.eridania; TISSUE=Head;  
RX MEDLINE=94176032; PubMed=8129851;  
RA Furuya K., Liao S., Reynolds S.E., Ota R.B., Hackett M.,  
RA Schooley D.A.;  
RT "Isolation and identification of a cardioactive peptide from Tenebrio  
RT molitor and Spodoptera eridania";  
RL Biol. Chem. Hoppe-Seyler 374:1065-1074(1993).  
CC -1- FUNCTION: THE EFFECT OF CCAP IS BOTH INO- AND CHRONOTROPIC.  
CC -1- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED  
CC INTO THE HEMOLYMPH.

DR PIR; A26363; A26363.  
DR PIR; S27233; S27233.  
KW Neuropeptide; Amidation.  
FT DISULFID 3 9  
FT MOD RES 9 9  
SQ SEQUENCE 9 AA, 959 MW; C5A861A9CD44EB9 CRC64;

## Query Match

Best Local Similarity 28.8%; Score 15; DB 1; Length 9;  
Matches 2; Conservativity 100.0%; Pred. No. 1.1e+05;  
Mismatches 0; Indels 0; Gaps 0;

4 FC 5  
||

Db 2 FC 3

RESULT 2

SAP\_STOVA STANDARD; PRT; 9 AA.

ID SAP\_STOVA

AC P24047;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DE 01-MAR-1992 (Rel. 21, Last annotation update)

OC Sperm-activating peptide (SAP).

OS Stomopneustes variolaris (Sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidae; Echinoidea; Diademata; Phymosomatoida; Stomechinidae;

OC Stomopneustes.

NCBI\_TaxID=7663;

RM NCBI\_TaxID=7663;

RP SEQUENCE, AND DISULFIDE BOND.

RC TISSUE=Egg jelly;

RX MEDLINE=92097763; PubMed=1756858;

RA Yoshino K.-I., Takao T., Shimonishi Y., Suzuki N.;

RT "determination of the amino acid sequence of an intramolecular disulfide linkage-containing sperm-activating peptide by tandem mass spectrometry."

RL FEBS Lett. 294:179-182(1991).

CC -1- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF CAMP, CGMP AND CLACIUM LEVELS IN SPERM CELLS. AND TRANSIENT ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF CC GUANYLATE CYCLASE.

CC PIR: S19329; S19329

DR DISULFID

FT SEQUENCE 9 AA; 1010 MW; C469B3387B076EB9 CRC64;

SQ

Query Match 28.8%; Score 15; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FC 5

Db 2 FC 3

RESULT 3

UPA3\_HUMAN STANDARD; PRT; 9 AA.

ID UPA3\_HUMAN

AC P30089;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Unknown protein from 2D-page of plasma (Spot 11) (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI\_TaxID=9606;

RM NCBI\_TaxID=9606;

RP SEQUENCE.

RC TISSUE=Plasma;

RX MEDLINE=93092937; PubMed=1459097;

RA Hughes G.J., Frutiger S., Paquet N., Pasquali C.,

RA Sanchez J.-C., James R., Tisot J.-D., Bjelqvist B.,

RA Hochstrasser D.F.;

RT "Plasma protein map: an update by microsequencing."

RL Electrophoresis 13:707-714(1992).

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 4.6, ITS MW IS: 46 kDa.

DR SWISS-2DPAGE; P30089; HUMAN.

FT NON TER

FT SEQUENCE 9 AA; 1056 MW; 26F2B1BAF769C737 CRC64;

SQ

Query Match 26.9%; Score 14; DB 1; Length 9;

Best Local Similarity 50.0%; Pred. No. 1.1e+05;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LFCASD 8

Db 3 LFPXTD 8

RESULT 4

FIBB\_CERSI STANDARD; PRT; 10 AA.

ID FIBB\_CERSI

AC P14537;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).

GN FGB.

OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.

NCBI\_TaxID=9807;

RM NCBI\_TaxID=9807;

RP SEQUENCE.

RA O'Reil P.B., Doolittle R.F.;

RT "Mammalian phylogeny based on fibrinopeptide amino acid sequences."

RL Syst. Zool. 22:590-595(1973).

CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET AGGREGATION.

CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.

CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR InterPro: IPR002181; Fibrinogen C.

DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.

KM Blood coagulation; Plasma.

FT PEPTIDE

FT NON TER

FT SEQUENCE 10 AA; 1097 MW; 9402B2BCDDDD33A CRC64;

SQ

Query Match 26.9%; Score 14; DB 1; Length 10;

Best Local Similarity 60.0%; Pred. No. 4.7e+03;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 ASDAK 10

Db 6 AVDAR 10

RESULT 5

URA6\_HUMAN STANDARD; PRT; 10 AA.

ID URA6\_HUMAN

AC P32080;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Unknown protein from 2D-page of red blood cells (Spot 17) (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI\_TaxID=9606;

RM NCBI\_TaxID=9606;

RP SEQUENCE.

RC TISSUE=Erythrocyte;

RX MEDLINE=94147970; PubMed=8313871;

RA Golaz O., Hughes G.J., Frutiger S., Paquet N., Bairoch A.,

RA Pasquali C., Sanchez J.-C., Tisot J.-D., Appel R.D., Walzer C.,

RA Bialant L., Hochstrasser D.F.;

RT "Plasma and red blood cell protein maps: update 1993."

RL Electrophoresis 14:1223-1231(1993).

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 5.4, ITS MW IS: 65 kDa.

DR SWISS-2DPAGE; P32080; HUMAN.  
 FT NON TER 10 10  
 SQ SEQUENCE 10 AA; 965 MW; 63DDC8086AE1EDDB CRC64;

Query Match 26.9%; Score 14; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 4.7e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 ASDA 9  
 |||  
 1 ASEA 4

## RESULT 6

UXA2\_CHLTR STANDARD; PRT; 10 AA.  
 ID \_UXA2\_CHLTR  
 AC P38003;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Unknown protein from 2D-page from elementary body (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 NC NCB1\_TaxID=813;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=12/434/Bu;  
 RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,  
 RA Comanducci M., Christensen G., Birkelund S., Vitreou E., Ratci G.,  
 RA Pallini V.;  
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.  
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 DR PROTEIN IS: 4.65, ITS MW IS: 26.5 KDa.  
 SIena-2DPAGE: P38003; -.  
 FT NON TER 10 10  
 SQ SEQUENCE 10 AA; 1019 MW; 49C285C6CAE862C7 CRC64;

Query Match 26.9%; Score 14; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 ASD 8  
 |||  
 Db 8 ASD 10

## RESULT 7

UPA9\_HUMAN STANDARD; PRT; 10 AA.  
 ID \_UPA9\_HUMAN  
 AC P30095;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 35) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing."  
 RL Electrophoresis 13:707-714(1992). THE DETERMINED PI OF THIS UNKNOWN  
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 DR PROTEIN IS: 7.2, ITS MW IS: 15 KDa.  
 SI SWISS-2DPAGE: P30095; HUMAN.  
 FT NON TER 10 10  
 SQ SEQUENCE 10 AA; 1233 MW; 37AD72B409C681B7 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 7.4e+03;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TLF 4  
 |||  
 Db 4 TMF 6

## RESULT 8

UC24\_MAIZE STANDARD; PRT; 7 AA.  
 ID \_UC24\_MAIZE  
 AC P80630;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 447)  
 DE (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 NC NCB1\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Coleoptile;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program."  
 RL Theor. Appl. Genet. 93:997-1006(1996).  
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 DR PROTEIN IS: 6.0, ITS MW IS: 30.0 KDa.  
 SI Maize-2DPAGE: P80630; COLEOPTILE.  
 FT NON TER 1 1  
 SQ SEQUENCE 7 AA; 665 MW; 6DC1B5B3DC1B5D0 CRC64;

Query Match 23.1%; Score 12; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 SDAK 10  
 |||  
 Db 1 STAK 4

## RESULT 9

ACT\_CARMA STANDARD; PRT; 8 AA.  
 ID \_ACT\_CARMA  
 AC P80709;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Actin (Fragment).  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 NC NCB1\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,  
 RA Baghdassarian D.;  
 RT "A transaldolase. An enzyme implicated in crab steroidogenesis."  
 RL Endocrine 5:23-32(1996).  
 CC -I- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED  
 CC IN ALL EUKARYOTIC CELLS.  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:

```

CC      6.8, ITS MM IS: 46 kDa.
CC      -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
DR      InterPro; IPR004001; Actin.
DR      InterPro; IPR004000; Actin-like.
DR      PROSITE; PS00406; ACTINS_1; PARTIAL.
DR      PROSITE; PS00432; ACTINS_2; PARTIAL.
DR      PROSITE; PS01132; ACTINS_ACT-LIKE; PARTIAL.
KM      Structural protein.
FT      NON_TER      1
FT      NON_TER      8
SQ      SEQUENCE      8 AA; 976 MW; 1424005AB2CAEB3 CRC64;

Query Match      23.1%; Score 12; DB 1; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 CASDAK 10
DB      2 CVDIR 7

RESULT 10
OMY_ORCLI      STANDARD; PRT; 8 AA.
AC      P82455;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Orcomyotopin (OMT).
OS      Orcomyotopin limosus (Spinycheek crayfish).
OC      Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC      Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC      Asacidea; Asacidea; Cambaridae; Orcomyotopin.
OX      NCBI_TaxID=28379;
RN      [1]
RP      SEQUENCE, MASS SPECTROMETRY, AND AMIDATION.
RC      TISSUE=Hindgut;
RA      MEDLINE=20411310; PubMed=10952880;
RA      Dirksen H., Burdzik S., Sauter A., Kellner R.;
RT      "Two orcomyotopins and the novel octapeptide orcomyotopin in the hindgut
RT      of the crayfish Orcomyotopin limosus: identified myostimulatory
RT      neuropeptides originating together in neurones of the terminal
RT      abdominal ganglion."
RL      J. Exp. Biol. 203:2807-2818(2000).
CC      -1- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
CC      AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS. IT IS SYNTHESIZED
CC      BY ABOOMINAL GANGLIONIC NEURONS.
CC      -1- MASS SPECTROMETRY: MM=904.8; METHOD=FAIR.
KM      Amidation; Neuropeptide.
FT      MOD_RES      8
FT      SEQUENCE      8 AA; 905 MW; 870861B1A9CDDAA9 CRC64;

Query Match      23.1%; Score 12; DB 1; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTIF 4
DB      5 TTGF 8

RESULT 11
COXA_ONCMY      STANDARD; PRT; 10 AA.
AC      P80328;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
OS      Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

```

```

OX      NCBI_TaxID=8022;
RN      [1]
RP      SEQUENCE.
RC      TISSUE=Liver;
RA      MEDLINE=94237150; PubMed=8181469;
RA      Freund R., Kadenbach B.;
RT      "Identification of tissue-specific isoforms for subunits Vb and VIIa
RT      of cytochrome c oxidase isolated from rainbow trout."
RL      Eur. J. Biochem. 221:1111-1116(1994).
CC      -1- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
CC      OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC      -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC      c + 2 H(2)O.
DR      PIR, S43625; S43625.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC      -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
KM      Oxidoreductase; Heme; Mitochondrion; Inner membrane.
FT      NON_TER      10
FT      NON_TER      10
SQ      SEQUENCE      10 AA; 1144 MW; C535C5B1AB02C33D CRC64;

Query Match      23.1%; Score 12; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      7 SDAK 10
DB      1 SHAK 4

RESULT 12
COXQ_SHEEP      STANDARD; PRT; 10 AA.
AC      P80337;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Cytochrome c oxidase polypeptide VIII-liver/heart (EC 1.9.3.1)
DE      (Fragment).
OS      Ovis aries (Sheep).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Caprinae; Ovis.
OX      NCBI_TaxID=9940;
RN      [1]
RP      SEQUENCE.
RC      TISSUE=Heart, and Liver;
RA      Freund R., Kadenbach B.;
RL      Submitted (MAR-1994) to the SWISS-PROT data bank.
CC      -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC      CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC      MITOCHONDRIAL ELECTRON TRANSPORT.
CC      -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC      c + 2 H(2)O.
CC      -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIII FAMILY.
KM      Oxidoreductase; Mitochondrion.
FT      NON_TER      10
FT      NON_TER      10
SQ      SEQUENCE      10 AA; 1027 MW; C4E95CA33DC7633D CRC64;

Query Match      23.1%; Score 12; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.1e+04;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      6 ASDAK 10
DB      3 AKPAK 7

RESULT 13
CCFI_ENTFA      STANDARD; PRT; 7 AA.
ID      CCFI_ENTFA
AC      P20104;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)

```

DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE Sex pheromone CCF10.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 NX NCB1\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=89008313; PubMed=3139658;  
 RA Mori M., Sakagami Y., Ishii Y., Isegai A., Kitada C., Fujino M.,  
 RA Adelt J.C., Dunny G.M., Suzuki A.;  
 RT "Structure of cCF10, a peptide sex pheromone which induces  
 RT conjugative transfer of the Streptococcus faecalis tetracycline  
 RT resistance plasmid, pCF10.";  
 RL J. Biol. Chem. 263:14574-14578(1988).  
 CC -I- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE  
 CC PIR: A30812; A30812.  
 DR Pheromone.  
 KW SEQUENCE  
 SQ SEQUENCE 7 AA; 790 MW; 72C9D2C71B2C740 CRC64;  
 Query Match 21.2%; Score 11; DB 1; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TTVF 4  
 DB 3 TLVF 6  
 RESULT 14  
 FAR8 CALVO STANDARD; PRT; 8 AA.  
 ID P41863;  
 AC 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CalliFMRamide 8.  
 OS Calliphora vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.  
 NX NCB1\_TaxID=27454;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Thoracic ganglion;  
 RC MEDLINE=92196111; PubMed=1549595;  
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
 RA Rehfeld J.F., Thorpe A.;  
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
 RT neuropeptides (designated calliFMRamides) from the blowfly  
 RT Calliphora vomitoria.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -I- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC PIR: H41978; H41978.  
 DR Neuropeptide; Amidation.  
 FT MOD RES 8  
 SQ SEQUENCE 8 AA; 957 MW; 72D40699CA44D8 CRC64;  
 Query Match 21.2%; Score 11; DB 1; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 ASD 8  
 DB 2 AND 4  
 RESULT 15  
 DSIP\_RABIT STANDARD; PRT; 9 AA.  
 ID DSIP\_RABIT  
 AC P01158;  
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Delta sleep-inducing peptide (DSIP).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NX NCB1\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77185324; PubMed=862769;  
 RA Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,  
 RA Schoenberger G.A.;  
 RT "The delta sleep inducing peptide (DSIP). Comparative properties of  
 RT the original and synthetic nonapeptide.";  
 RL Experientia 33:548-552(1977).  
 RN [2]  
 RP SEQUENCE, AND SYNTHESIS.  
 RX MEDLINE=79054421; PubMed=568769;  
 RA Schoenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.,  
 RT "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid  
 RT analysis, sequence, synthesis and activity of the nonapeptide.";  
 RL Pflügers Arch. 376:119-129(1978).  
 RN [3]  
 RP REVIEW.  
 RX MEDLINE=87175129; PubMed=3550726;  
 RA Graf M.V., Kastin A.J.;  
 RL "Delta-sleep-inducing peptide (DSIP): an update.";  
 CC Peptides 7:1165-1187(1986).  
 CC -I- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF  
 CC RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND  
 CC REDUCED MOTOR ACTIVITIES.  
 CC -I- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF  
 CC OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC  
 CC STIMULATION OF THE THALAMUS.  
 CC -I- DATABASE: NAME=Protein SpotLight;  
 CC NOTE=Issue 8 of March 2001;  
 CC WWW="http://www.expaasy.org/spotlight/articles/sp11c008.html".  
 DR PIR: A01422; ODRB.  
 SQ SEQUENCE 9 AA; 849 MW; DDD365BDDA8787D CRC64;  
 Query Match 21.2%; Score 11; DB 1; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 1.1e+05;  
 Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 FCASDA 9  
 DB 1 WAGSDA 6

Search completed: April 14, 2003, 07:03:12  
 Job time : 13 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 06:43:27 ; Search time 29 seconds

(without alignments)  
71.051 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTTFPCASDAK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1224

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_21:\*  
2: sp archaea:\*  
3: sp bacteria:\*  
4: sp fungi:\*  
5: sp human:\*  
6: sp invertebrate:\*  
7: sp mammal:\*  
8: sp organelle:\*  
9: sp phage:\*  
10: sp plant:\*  
11: sp rodent:\*  
12: sp virus:\*  
13: sp vertebrate:\*  
14: sp unclassified:\*  
15: sp virus:\*  
16: sp bacteriophage:\*  
17: sp archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	38.5	9	12 O90350	O90350 hepatitis g
2	19	36.5	7	11 O55184	O55184 rattus norv
3	19	36.5	8	10 Q98AY7	Q98AY7 dioscorea t
4	19	36.5	10	2 Q44451	Q44451 agrobacteri
5	17	32.7	10	2 Q9K343	Q9K343 escherichia
6	17	32.7	10	4 Q9UN90	Q9UN90 homo sapien
7	16	30.8	8	2 Q8RSR3	Q8RSR3 lactobacill
8	16	30.8	8	3 Q9URB9	Q9URB9 saccharomyc
9	16	30.8	8	4 Q9BYV5	Q9BYV5 homo sapien
10	16	30.8	8	6 Q9BFC3	Q9BFC3 didelphis m
11	16	30.8	8	6 Q9BFC2	Q9BFC2 macropus eu
12	16	30.8	8	6 Q9BFC1	Q9BFC1 choleopus d
13	16	30.8	8	6 Q9BFC0	Q9BFC0 choleopus d
14	16	30.8	8	6 Q9BFB9	Q9BFB9 euphractus
15	16	30.8	8	6 Q9BFB8	Q9BFB8 chaetophrac
16	16	30.8	8	6 Q9BFB7	Q9BFB7 tamandua te

17	16	30.8	8	6 Q9BFB6	Q9BFB6 myrmecophag
18	16	30.8	8	6 Q9BFB5	Q9BFB5 etinaceus c
19	16	30.8	8	6 Q9BFB4	Q9BFB4 talpa alta
20	16	30.8	8	6 Q9BFB3	Q9BFB3 condylura c
21	16	30.8	8	6 Q9BFB2	Q9BFB2 sorex arane
22	16	30.8	8	6 Q9BFB1	Q9BFB1 echinops te
23	16	30.8	8	6 Q9BFB0	Q9BFB0 trichechus
24	16	30.8	8	6 Q9BFA9	Q9BFA9 procavia ca
25	16	30.8	8	6 Q9BFA8	Q9BFA8 toxodonta a
26	16	30.8	8	6 Q9BFA6	Q9BFA6 cyroteropus
27	16	30.8	8	6 Q9BFA5	Q9BFA5 cynocephalu
28	16	30.8	8	6 Q9BFA4	Q9BFA4 tupia mino
29	16	30.8	8	6 Q9BFA3	Q9BFA3 lemur catta
30	16	30.8	8	6 Q9BFA2	Q9BFA2 tarsius ban
31	16	30.8	8	6 Q9BFA1	Q9BFA1 ateles fusc
32	16	30.8	8	6 Q9BFA0	Q9BFA0 macaca mla
33	16	30.8	8	6 Q9BF99	Q9BF99 hylobates c
34	16	30.8	8	6 Q9BF98	Q9BF98 callimico g
35	16	30.8	8	6 Q9BF97	Q9BF97 artibeus ja
36	16	30.8	8	6 Q9BF96	Q9BF96 pretopus gi
37	16	30.8	8	6 Q9BF95	Q9BF95 rousettus l
38	16	30.8	8	6 Q9BF94	Q9BF94 nycteris ch
39	16	30.8	8	6 Q9BF93	Q9BF93 megaptera n
40	16	30.8	8	6 Q9BF92	Q9BF92 tursiops tr
41	16	30.8	8	6 Q9BF91	Q9BF91 hippopotamu
42	16	30.8	8	6 Q9BF90	Q9BF90 tragelaphus
43	16	30.8	8	6 Q9BF89	Q9BF89 okapia john
44	16	30.8	8	6 Q9BF88	Q9BF88 equus cabal
45	16	30.8	8	6 Q9BF87	Q9BF87 tapirus ind

## ALIGNMENTS

RESULT 1  
ID O90350 PRELIMINARY; PRT; 9 AA.  
AC O90350;  
DT 01-NOV-1998 (T-EMBLrel. 08, Created)  
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis G virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC GBV-C/HGV group.  
OX NCBI\_TaxID=45255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SG3419;  
RX MEDLINE=99266893; PubMed=10335862;  
RA Wong S.B.J., Chan S.H., Ren E.C.;  
RT "Diversity of GB virus C/hepatitis G virus isolates in Singapore:  
predominance of group 2a and the Asian group 3 variant.";  
RL J. Med. Virol. 58:145-153(1999).  
FR EMBL; AF078065; AAC32371.1; -.  
DT NON TER  
SQ SEQUENCE 9 AA; 989 MW; D9SCA5A5BB9CDD CRC64;

Query Match 38.5%; Score 20; DB 12; Length 9;  
Best Local Similarity 75.0%; Pred. No. 6.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 FCAS 7  
Db 4 FCSS 7

RESULT 2  
ID O55184 PRELIMINARY; PRT; 7 AA.  
AC O55184;  
DT 01-JUN-1998 (T-EMBLrel. 06, Created)  
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Orphan receptor TR4-NS (Fragment).  
 GN TR4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=96198747; PubMed=8612486;  
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,  
 RA Decera-Wadleigh S.D.;  
 RT "Splice variants of rat TR4 orphan receptor: differential expression  
 RT of novel sequences in the 5'-untranslated region and C-terminal  
 RT domain.";  
 RL Endocrinology 137:1562-1571(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=96299786; PubMed=8661150;  
 RA Yoshikawa T., Dupont B.R., Leach R.J., Decera-Wadleigh S.D.;  
 RT "New variants of the human and rat nuclear hormone receptor, TR4:  
 RT expression and chromosomal localization of the human gene.";  
 RL Genomics 35:361-366(1996).  
 DR EMBL; U59454; AAB91433.1; -.  
 KW Receptor.  
 FT NON\_TER  
 SQ SEQUENCE 7 AA; 663 MW; 6DDA8787B05350 CRC64;

Query Match 36.5%; Score 19; DB 11; Length 7;  
 Best Local Similarity 60.0%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 CASDA 9  
 DB 3 CGSDA 7

RESULT 3  
 OQSAY7 PRELIMINARY; PRT; 8 AA.  
 ID OQSAY7  
 AC OQSAY7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Phosphoglucose isomerase (Fragment).  
 OS Diocotyle tokoro.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;  
 OC Dioscorea.  
 OX NCBI\_TaxID=64475;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DT17-1.  
 RX MEDLINE=20279211; PubMed=10821191;  
 RA Terauchi R., Kahl G.;  
 RT "Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking  
 RT regions of Pal and Pgi genes from yams (Dioscorea).";  
 RL Mol. Gen. Genet. 263:554-560(2000).  
 DR EMBL; AB016716; BAA32235.1; -.  
 KW Isomerase.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 839 MW; FTB05731B5A1ADD6 CRC64;

Query Match 36.5%; Score 19; DB 10; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TLFC 5  
 DB 1 STLIC 8

RESULT 4  
 ID O44451 PRELIMINARY; PRT; 10 AA.  
 AC O44451;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Ti-plasmid DNA for nopaline synthetase promoter (Fragment).  
 OS Agrobacterium tumefaciens.  
 OG Plasmid T1.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85037947; PubMed=6493982;  
 RA Shaw C.H., Carter G.H., Watson M.D., Shaw C.H.;  
 RT "A functional map of the nopaline synthase promoter.";  
 RL Nucleic Acids Res. 12:7831-7846(1984).  
 DR EMBL; X01077; CAA25553.1; -.  
 KW Plasmid.  
 FT NON\_TER  
 SQ SEQUENCE 10 AA; 1067 MW; 1A6F98ADD9C731A0 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TLFCAS 7  
 DB 4 TLFCAS 9

RESULT 5  
 ID O9K343 PRELIMINARY; PRT; 10 AA.  
 AC O9K343;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Beta-lactamase (Fragment).  
 GN AMPC.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E3, E1, E2, E4, AND E5;  
 RX MEDLINE=99216914; PubMed=10103209;  
 RA Nelson E.C., Elisha B.G.;  
 RT "Molecular basis of AmpC hyperproduction in clinical isolates of  
 RT Escherichia coli.";  
 RL Antimicrob. Agents Chemother. 43:957-959(1999).  
 DR EMBL; AF119773; AAF28860.1; -.  
 DR EMBL; AF119769; AAF28856.1; -.  
 DR EMBL; AF119770; AAF28857.1; -.  
 DR EMBL; AF119771; AAF28858.1; -.  
 DR EMBL; AF119772; AAF28859.1; -.  
 DR InterPro; IPR001230; Premyl site.  
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN\_1.  
 FT NON\_TER  
 SQ SEQUENCE 10 AA; 1140 MW; 46A614DEB731B1A3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 6.7e+03;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TLFC 6  
 DB 4 TLFC 8

RESULT 6  
Q9UN90 PRELIMINARY; PRT; 10 AA.  
AC Q9UN90;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE Canalicular multispecific organic anion transporter (Fragment).  
GN CMOAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tanaka T., Uchimi T., Hinoshita E., Inokuchi A., Toh S., Wada M.,  
RA Nomoto M., Kohno K., Kuwano M.;  
RT "Sequence analysis and functional characterization of the 5'-flanking  
RT region of the human canalicular multispecific organic anion  
RT transporter/multidrug resistance protein 2 (CMOAT/MRP2) gene.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF144630; AAD47599.1; -.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1219 MW; 76F28CB4EB9C3B CRC64;

Query Match 32.7%; Score 17; DB 4; Length 10;  
Best Local Similarity 75.0%; Pred. No. 6.7e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 FCAS 7  
DB 5 FCNS 8

RESULT 7  
Q8RSR3 PRELIMINARY; PRT; 8 AA.  
AC Q8RSR3;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE PduG protein (Fragment).  
GN PDUg.  
OS Lactobacillus collinoides.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Lactobacillaceae; Lactobacillus.  
OX NCBI\_TaxID=33960;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=LMG 18850;  
RA Sauvaget N., Gouffé K., Laplace J.M., Auffray Y.;  
RT "Characterisation of the diol dehydratase pdu operon of Lactobacillus  
RT collinoides.";  
RL FEMS Microbiol. Lett. 0:0-0(0).  
DR EMBL; AJ297723; CAD01093.1; -.  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 882 MW; ECA40B05BABSBD6 CRC64;

Query Match 30.8%; Score 16; DB 2; Length 8;  
Best Local Similarity 60.0%; Pred. No. 6.7e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 ASDAK 10  
DB 2 ASDSE 6

RESULT 8  
Q9URB9 PRELIMINARY; PRT; 8 AA.  
AC Q9URB9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Aminopeptidase B (Fragment).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE  
RA MEDLINE=92088139; PubMed=1750699;  
RA Kassel D.B., Williams K.P., Musselman B.D., Smith J.A.;  
RT "Optimization of the fragmentation in a first-fast atom bombardment ion  
RT source for the sequencing of peptides at the picomole level.";  
RL Anal. Chem. 63:1978-1983(1991).  
FT NON\_TER 1 1  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 772 MW; 783DDAADCC2C732C8 CRC64;

Query Match 30.8%; Score 16; DB 3; Length 8;  
Best Local Similarity 75.0%; Pred. No. 6.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 SDAK 10  
DB 5 ADAX 8

RESULT 9  
Q9BY5 PRELIMINARY; PRT; 8 AA.  
AC Q9BY5;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE CAMP responsive element modulator (Fragment).  
GN CREM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21082082; PubMed=11214319;  
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
RA O'Brien S.J.;  
RT "Molecular phylogenetics and the origins of placental mammals.";  
RL Nature 409:614-618(2001).  
DR EMBL; AY011664; AAG47575.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EFAB572A CRC64;

Query Match 30.8%; Score 16; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LFC 5  
DB 2 LVC 4

RESULT 10  
Q9BFC3 PRELIMINARY; PRT; 8 AA.  
AC Q9BFC3;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE CAMP responsive element modulator (Fragment).  
GN CREM.  
OS Didelphis marsupialis virginiana (North American opossum).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
OX NCBI\_TaxID=9267;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001);
DR EMBL: AY011620; AAC47535.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 978 MW; DF1DD31EAB572A CRC64;

Query Match
Best Local Similarity 30.8%; Score 16; DB 6; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 LFC 5
Db 2 LFC 4

RESULT 11
Q9BFC2 PRELIMINARY; PRT; 8 AA.
AC Q9BFC2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001);
DR EMBL: AY011621; AAC47536.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 978 MW; DF1DD31EAB572A CRC64;

Query Match
Best Local Similarity 30.8%; Score 16; DB 6; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 LFC 5
Db 2 LFC 4

RESULT 12
Q9BFC1 PRELIMINARY; PRT; 8 AA.
AC Q9BFC1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Cholepus hoffmanni (Hoffmann's two-fingered sloth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Edentata; Cholepidae; Cholepus.
OX NCBI_TaxID=9358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001);

```

```

DR EMBL: AY011622; AAC47537.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 978 MW; DF1DD31EAB572A CRC64;

Query Match
Best Local Similarity 30.8%; Score 16; DB 6; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 LFC 5
Db 2 LFC 4

RESULT 13
Q9BFC0 PRELIMINARY; PRT; 8 AA.
AC Q9BFC0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Cholepus didactylus (southern two-toed sloth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Edentata; Cholepidae; Cholepus.
OX NCBI_TaxID=27675;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001);
DR EMBL: AY011623; AAC47538.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 978 MW; DF1DD31EAB572A CRC64;

Query Match
Best Local Similarity 30.8%; Score 16; DB 6; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 LFC 5
Db 2 LFC 4

RESULT 14
Q9BFB9 PRELIMINARY; PRT; 8 AA.
AC Q9BFB9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Euphractus sexcinctus (Six-banded armadillo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Edentata; Dasypodidae; Euphractus.
OX NCBI_TaxID=143300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001);
DR EMBL: AY011624; AAC47539.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 978 MW; DF1DD31EAB572A CRC64;

Query Match
Best Local Similarity 30.8%; Score 16; DB 6; Length 8;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 LFC 5  
 Db 2 LFC 4

## RESULT 15

O9BFB8  
 ID O9BFB8 PRELIMINARY; PRT; 8 AA.  
 AC O9BFB8;  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE CAMP responsive element moderator (Fragment).  
 GN CREM.  
 OS Chaetophractus villosus (South American armadillo).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Edentata; Daeypodidae; Chaetophractus.  
 OX NCBI\_TaxID=29080;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals.";  
 RL Nature 409:614-618(2001).  
 DR EMBL; AY011625; AAC47540.1; -.  
 FT NON TER 1  
 SQ SEQUENCE 8 AA; 978 MW; DF1DD331EAB572A CRC64;

Query Match 30.8%; Score 16; DB 6; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LFC 5  
 Db 2 LFC 4

Search completed: April 14, 2003, 07:05:05  
 Job time : 30 secs

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 07:05:12 ; Search time 14 Seconds

(without alignments)  
43.669 Million cell updates/sec

Title: US-09-017-524A-32

Perfect score: 52

Sequence: 1 TTFFCASDAK 10

Scoring table: BIOSIM62

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 29968

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB\_PEP:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB\_PEP:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB\_PEP:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB\_PEP:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB\_PEP:\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB\_PEP:\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB\_PEP:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB\_PEP:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB\_PEP:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB\_PEP:\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB\_PEP:\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB\_PEP:\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB\_PEP:\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB\_PEP:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	1	US-08-821-739A-89
2	52	100.0	10	1	US-09-894-018-164
3	47	90.4	9	1	US-08-821-739A-67
4	38	73.1	7	10	US-09-911-838-53
5	38	73.1	7	10	US-09-911-838-57
6	37	71.2	7	10	US-09-911-838-51
7	37	71.2	7	10	US-09-911-838-55
8	33	63.5	6	10	US-09-911-838-50
9	33	63.5	6	10	US-09-911-838-56
10	33	63.5	6	10	US-09-911-838-49
11	33	63.5	6	10	US-09-911-838-54
12	32	61.5	6	10	US-09-911-838-58
13	32	61.5	6	10	US-09-911-838-52
14	32	61.5	6	10	US-09-911-838-56
15	29	55.8	6	10	US-09-909-950-23
16	29	55.8	6	10	US-09-911-838-48
17	29	55.8	6	10	US-09-911-838-47
18	28	53.8	9	10	US-09-834-765-216
19	28	53.8	9	10	US-09-834-765-321

20	28	53.8	10	10	US-09-834-765-307	Sequence 307, App
21	28	53.8	10	10	US-09-834-765-374	Sequence 374, App
22	24	46.2	9	10	US-09-834-765-57	Sequence 57, App
23	24	46.2	10	10	US-09-834-765-113	Sequence 113, App
24	23	44.2	6	10	US-09-911-838-60	Sequence 60, App
25	23	44.2	7	10	US-09-911-838-61	Sequence 61, App
26	23	44.2	8	9	US-09-910-552-15	Sequence 15, App
27	23	44.2	8	10	US-09-789-697A-14	Sequence 14, App
28	23	44.2	10	9	US-10-097-065-402	Sequence 402, App
29	22	42.3	9	9	US-09-930-559-5	Sequence 5, App
30	22	42.3	9	10	US-09-894-018-224	Sequence 224, App
31	21	40.4	8	9	US-09-758-426-51	Sequence 51, App
32	21	40.4	8	9	US-09-945-249-23	Sequence 23, App
33	21	40.4	8	9	US-09-945-249-28	Sequence 28, App
34	21	40.4	8	9	US-09-758-198-51	Sequence 51, App
35	21	40.4	8	9	US-09-861-661-51	Sequence 51, App
36	21	40.4	8	10	US-09-758-128-51	Sequence 51, App
37	21	40.4	8	10	US-09-359-325A-14	Sequence 14, App
38	21	40.4	8	10	US-09-359-325A-15	Sequence 15, App
39	21	40.4	9	9	US-10-102-283-134	Sequence 134, App
40	21	40.4	9	10	US-09-765-086-167	Sequence 167, App
41	21	40.4	9	10	US-09-919-048-134	Sequence 134, App
42	21	40.4	10	10	US-09-834-765-291	Sequence 291, App
43	21	40.4	10	10	US-09-834-765-376	Sequence 376, App
44	20	38.5	6	10	US-09-911-838-46	Sequence 46, App
45	20	38.5	7	10	US-09-911-838-45	Sequence 45, App

## ALIGNMENTS

RESULT 1  
US-08-821-739A-89  
Sequence 89, Application US/08821739A  
Patent No. US20020168174A1  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding Peptides and Their Uses  
FILE REFERENCE: 2060.005000A  
CURRENT FILING DATE: US/08/821,739A  
CURRENT FILING DATE: 1999-03-20  
PRIORITY APPLICATION NUMBER: 60/013,833  
PRIORITY FILING DATE: 1996-03-21  
PRIORITY APPLICATION NUMBER: 08/589,107  
PRIORITY FILING DATE: 1996-07-12  
PRIORITY APPLICATION NUMBER: 08/451,913  
PRIORITY FILING DATE: 1995-05-26  
PRIORITY APPLICATION NUMBER: 08/347,610  
PRIORITY FILING DATE: 1994-12-01  
PRIORITY APPLICATION NUMBER: 08/186,266  
PRIORITY FILING DATE: 1994-01-25  
PRIORITY APPLICATION NUMBER: 08/159,339  
PRIORITY FILING DATE: 1993-11-29  
PRIORITY APPLICATION NUMBER: 08/103,396  
PRIORITY FILING DATE: 1993-08-06  
PRIORITY APPLICATION NUMBER: 08/027,746  
PRIORITY FILING DATE: 1993-03-05  
PRIORITY APPLICATION NUMBER: 07/926,666  
PRIORITY FILING DATE: 1992-08-07  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 89  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-821-739A-89

Query Match 100.0%; Score 52; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TILFCASDAK 10  
|||||  
DB 1 TILFCASDAK 10

## RESULT 2

US-09-894-018-164  
Sequence 164, Application US/09894018  
Patent No. US20020119127A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE, Inc.  
APPLICANT: Sette, Alessandro  
APPLICANT: Chestnut, Robert  
APPLICANT: Livingston, Brian  
APPLICANT: Baker, Denlaw  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David  
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894,018  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/284,221  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 164  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Transgenic mouse  
US-09-894-018-164

Query Match 100.0%; Score 52; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TILFCASDAK 10  
|||||  
DB 1 TILFCASDAK 10

## RESULT 3

US-08-821-739A-67  
Sequence 67, Application US/08821739A  
Patent No. US20020168374A1  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esben  
TITLE OF INVENTION: HLA Binding Peptides and Their Uses  
FILE REFERENCE: 2060.005000A  
CURRENT APPLICATION NUMBER: US/08/821,739A  
CURRENT FILING DATE: 1999-03-20  
PRIOR APPLICATION NUMBER: 60/013,833  
PRIOR FILING DATE: 1996-03-21  
PRIOR APPLICATION NUMBER: 08/589,107  
PRIOR FILING DATE: 1996-07-12  
PRIOR APPLICATION NUMBER: 08/451,913  
PRIOR FILING DATE: 1995-05-26  
PRIOR APPLICATION NUMBER: 08/347,610  
PRIOR FILING DATE: 1994-12-01  
PRIOR APPLICATION NUMBER: 08/186,266  
PRIOR FILING DATE: 1994-01-25  
PRIOR APPLICATION NUMBER: 08/159,339  
PRIOR FILING DATE: 1993-11-29  
PRIOR APPLICATION NUMBER: 08/103,396  
PRIOR FILING DATE: 1993-08-06

PRIOR APPLICATION NUMBER: 08/027,746  
PRIOR FILING DATE: 1993-03-05  
PRIOR APPLICATION NUMBER: 07/926,666  
PRIOR FILING DATE: 1992-08-07  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 67  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-821-739A-67

Query Match 90.4%; Score 47; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TILFCASDAK 10  
|||||  
DB 1 TILFCASDAK 9

## RESULT 4

US-09-911-838-53  
Sequence 53, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPRYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
TITLE OF INVENTION: SYNDROME  
FILE REFERENCE: UTSC:267USCI  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1993-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 53  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-911-838-53

Query Match 73.1%; Score 38; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TILFCASD 8  
|||||  
DB 1 TILFCASD 7

## RESULT 5

US-09-911-838-57  
Sequence 57, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPRYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
TITLE OF INVENTION: SYNDROME  
FILE REFERENCE: UTSC:267USCI  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1992-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 57  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-911-838-57

Query Match 73.1%; Score 38; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFCASDAK 10  
DB 1 LFCASDAK 7

RESULT 6  
US-09-911-838-51  
Sequence 51, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
FILE REFERENCE: UTSC:267USC1  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1992-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 51  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-51

Query Match 71.2%; Score 37; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFAS 7  
DB 1 TTLCFAS 7

RESULT 7  
US-09-911-838-55  
Sequence 55, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
FILE REFERENCE: UTSC:267USC1  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1992-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 55  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-55

Query Match 71.2%; Score 37; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LFCASDA 9  
DB 1 LFCASDA 7

RESULT 8  
US-09-911-838-50  
Sequence 50, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
FILE REFERENCE: UTSC:267USC1  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1992-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 50  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-50

Query Match 63.5%; Score 33; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLCFA 6  
DB 1 TTLCFA 6

RESULT 9  
US-09-911-838-54  
Sequence 54, Application US/09911838  
Patent No. US20020151678A1  
GENERAL INFORMATION:  
APPLICANT: ARLINGHAUS, RALPH  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
FILE REFERENCE: UTSC:267USC1  
CURRENT APPLICATION NUMBER: US/09/911,838  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: 07/834,923  
PRIOR FILING DATE: 1992-02-13  
NUMBER OF SEQ ID NOS: 226  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 54  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-54

Query Match 63.5%; Score 33; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LFCASD 8  
DB 1 LFCASD 6

RESULT 10  
US-09-911-838-56  
; Sequence 56, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-56

Query Match 63.5%; Score 33; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FCASDA 9  
Db 1 FCASDA 6

RESULT 11  
US-09-911-838-49  
; Sequence 49, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-49

Query Match 63.5%; Score 33; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTLFCA 6  
Db 2 TTLFCA 7

RESULT 12  
US-09-911-838-52  
; Sequence 52, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY

; TITLE OF INVENTION: SYNDROME  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 52  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-52

Query Match 61.5%; Score 32; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLFCAS 7  
Db 1 TLFCAS 6

RESULT 13  
US-09-911-838-58  
; Sequence 58, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 58  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-911-838-58

Query Match 61.5%; Score 32; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CASDAK 10  
Db 1 CASDAK 6

RESULT 14  
US-09-911-838-59  
; Sequence 59, Application US/09911838  
; Patent No. US20020151678A1  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; FILE REFERENCE: UTSC:267USC1  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226

SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; US-09-911-838-59

Query Match 61.5%; Score 32; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 CASDAK 10  
Db 1 CASDAK 6

RESULT 15  
US-09-909-950-23  
; Sequence 23, Application US/09909950  
; Patent No. US2002011299A1  
; GENERAL INFORMATION:  
; APPLICANT: KOLBE, Hanno V.J.  
; RASMUSSEN, Ulla B.  
; KREIL, Gunther  
; ACHSTETTER, Tilman  
; TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/909,950  
; FILING DATE: 23-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/909,950  
; FILING DATE: 2001-01-23  
; APPLICATION NUMBER: FR 90 07901  
; FILING DATE: 29-JUN-1993  
; APPLICATION NUMBER: FR 94 00202  
; FILING DATE: 11-JAN-1994  
; APPLICATION NUMBER: FR 9400062.9  
; FILING DATE: 11-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meuth, Donna M.  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 017753-071  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..6  
; OTHER INFORMATION: /note="xenoxin-3, alkylated"

fragment T4"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
; US-09-909-950-23

Query Match 55.8%; Score 29; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FCASD 8  
Db 1 FCASD 5

Search completed: April 14, 2003, 07:13:03  
Job time : 15 secs

2

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2003, 07:03:00 ; Search time 27 Seconds  
(Without alignments)  
10.897 Million cell updates/sec

Title: US-09-017-524A-32  
Perfect score: 52  
Sequence: 1 TITLFCASDAK 10

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 77191

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	63.5	6	3	US-08-167-867-18	Sequence 18, Appl
2	33	63.5	6	4	US-09-317-993-18	Sequence 18, Appl
3	33	63.5	10	1	US-08-306-116A-42	Sequence 42, Appl
4	29	55.8	6	3	US-08-578-674-23	Sequence 23, Appl
5	29	55.8	6	4	US-09-498-346-23	Sequence 23, Appl
6	25	48.1	8	1	US-07-958-903A-51	Sequence 51, Appl
7	25	48.1	8	1	US-08-462-018-51	Sequence 51, Appl
8	25	48.1	8	1	US-08-823-245-51	Sequence 51, Appl
9	25	48.1	8	4	US-07-963-329A-51	Sequence 51, Appl
10	25	48.1	9	4	PCT-US92-09443A-51	Sequence 51, Appl
11	25	48.1	9	4	US-08-866-545-19	Sequence 19, Appl
12	25	48.1	10	1	US-07-958-903A-51	Sequence 19, Appl
13	25	48.1	10	1	US-08-462-018-51	Sequence 2, Appl
14	25	48.1	10	1	US-08-823-245-51	Sequence 2, Appl
15	25	48.1	10	1	US-07-963-329A-68	Sequence 68, Appl
16	25	48.1	10	4	US-07-963-329A-75	Sequence 68, Appl
17	25	48.1	10	5	PCT-US92-09443A-68	Sequence 75, Appl
18	25	48.1	10	5	PCT-US92-09443A-75	Sequence 75, Appl
19	24	46.2	8	1	US-08-526-710-11	Sequence 11, Appl
20	24	46.2	8	3	US-08-862-855-11	Sequence 11, Appl
21	24	46.2	8	4	US-09-226-985-11	Sequence 11, Appl
22	24	46.2	8	4	US-09-227-906-11	Sequence 11, Appl
23	23	44.2	9	3	US-08-925-002-15	Sequence 15, Appl
24	23	44.2	9	1	US-08-584-226-17	Sequence 17, Appl
25	22	42.3	8	1	US-07-958-903A-53	Sequence 53, Appl
26	22	42.3	8	1	US-08-462-018-53	Sequence 53, Appl
27	22	42.3	8	1	US-08-823-245-53	Sequence 53, Appl

28	22	42.3	8	4	US-07-963-329A-53	Sequence 53, Appl
29	22	42.3	8	5	PCT-US92-09443A-53	Sequence 53, Appl
30	22	42.3	9	3	US-08-159-339A-491	Sequence 491, Appl
31	22	42.3	9	4	US-08-866-545-20	Sequence 20, Appl
32	22	42.3	10	1	US-07-958-903A-46	Sequence 46, Appl
33	22	42.3	10	1	US-08-462-018-46	Sequence 46, Appl
34	22	42.3	10	1	US-08-823-245-46	Sequence 46, Appl
35	22	42.3	10	3	US-08-159-339A-501	Sequence 501, Appl
36	22	42.3	10	4	US-07-963-329A-46	Sequence 46, Appl
37	22	42.3	10	5	PCT-US92-09443A-46	Sequence 46, Appl
38	21	40.4	6	1	US-08-290-919-18	Sequence 18, Appl
39	21	40.4	8	4	US-08-842-306B-37	Sequence 37, Appl
40	21	40.4	8	4	US-08-842-306B-38	Sequence 37, Appl
41	21	40.4	8	4	US-08-838-973B-37	Sequence 37, Appl
42	21	40.4	8	4	US-08-838-973B-38	Sequence 38, Appl
43	21	40.4	8	4	US-08-771-212A-33	Sequence 33, Appl
44	21	40.4	8	4	US-08-771-212A-34	Sequence 34, Appl
45	21	40.4	9	3	US-08-159-339A-1154	Sequence 1154, Ap

## ALIGNMENTS

RESULT 1  
US-08-167-867-18  
Sequence 18, Application US/08167867  
Patent No. 6093405  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: IMMUNOGEN COMPOUNDS HAVING  
SPECIFICALLY AN ANTI-CYTOKINE EFFECT, METHOD OF  
PREPARATION, PHARMACEUTICAL COMPOSITIONS AND KITS  
TITLE OF INVENTION: CONTAINING THEM  
NUMBER OF SEQUENCES: 19  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,867  
FILING DATE: 17-DEC-1993  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-167-867-18  
Query Match 63.5%; Score 33; DB 3; Length 6;  
Best local Similarity 100.0%; Pred. No. 2e+05; Mismatches 0; Gaps 0;  
Matches 6; Conservative 0; Indels 0;  
Cy 1 TITLFC A 6  
Db 1 TITLFC A 6  
RESULT 2  
US-09-317-993-18  
Sequence 18, Application US/09317993  
Patent No. 6455045  
GENERAL INFORMATION:  
APPLICANT: ZAGURY Daniel  
TITLE OF INVENTION: IMMUNOGEN COMPOUNDS HAVING  
SPECIFICALLY AN ANTI-CYTOKINE EFFECT, METHOD OF  
PREPARATION, PHARMACEUTICAL COMPOSITIONS AND KITS  
CONTAINING THEM  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
STREET: 19 rue d'Eiancourt, La Verriere

CITY: le Mesnil-Saint-Denis  
COUNTRY: FRANCE  
ZIP: 78320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (ERO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/317,993  
FILING DATE: 25-May-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,867  
FILING DATE: 17-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BIZZINI Bernard  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-317-993-18

Query Match 63.5%; Score 33; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLFCA 6  
Db 1 TTLFCA 6

RESULT 3  
US-08-306-116A-42  
Sequence 42, Application US/08306116A  
Patent No. 5691135  
GENERAL INFORMATION:  
APPLICANT: Braun, Jonathan  
APPLICANT: Goodlick, Lee A.  
TITLE OF INVENTION: IMMUNOGLOBULIN SUPERANTIGEN BINDING TO GP120  
TITLE OF INVENTION: FROM HIV  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Knobbe, Marens, Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,116A  
FILING DATE: 14-SEP-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/259,669  
FILING DATE: 14-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Altman, Daniel E.  
REGISTRATION NUMBER: 34,115  
REFERENCE/DOCKET NUMBER: DCLA004,001CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 714-760-0404  
TELEFAX: 714-760-9502  
TELEX:  
INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-306-116A-42

Query Match 63.5%; Score 33; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTLFCA 6  
Db 5 TTLFCA 10

RESULT 4  
US-08-578-674-23  
Sequence 23, Application US/08578674  
Patent No. 6077827  
GENERAL INFORMATION:  
APPLICANT: Kolbe, Hanno V.J.  
APPLICANT: RASMUSSEN, Ulla B.  
APPLICANT: KREIL, Gunther  
APPLICANT: ACHSTERFER, Tilman  
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/578,674  
FILING DATE: 28-DEC-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 90 07901  
FILING DATE: 29-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 94 00302  
FILING DATE: 11-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9400062.9  
FILING DATE: 11-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 017753-071  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide

LOCATION: 1..6  
OTHER INFORMATION: /note="xenoxin-3, alkylated"  
OTHER INFORMATION: fragment T4"  
US-08-578-674-23

Query Match 55.8%; Score 29; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FCASD 8  
|||||  
Db 1 FCASD 5

RESULT 5  
US-09-498-346-23  
Sequence 23, Application US/09498346  
Patent No. 6277822  
GENERAL INFORMATION:  
APPLICANT: KOLBE, Hanno V.J.  
APPLICANT: RASMUSSEN, Ulla B.  
APPLICANT: KREIL, Gunther  
APPLICANT: ACHSTERER, Tilman  
TITLE OF INVENTION: FAMILY OF PEPTIDES KNOWN AS XENOXINS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/498,346  
FILING DATE: 04-FEB-2000  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/578,674  
FILING DATE: 28-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 90 07901  
FILING DATE: 29-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 94 00202  
FILING DATE: 11-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9400652.9  
FILING DATE: 11-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 017753-071  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURES:  
NAME/KEY: Peptide  
LOCATION: 1..6  
OTHER INFORMATION: /note="xenoxin-3, alkylated"  
OTHER INFORMATION: fragment T4"  
US-09-498-346-23

Query Match 55.8%; Score 29; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FCASD 8  
|||||  
Db 1 FCASD 5

RESULT 6  
US-07-958-903A-51  
Sequence 51, Application US/07958903A  
Patent No. 5652214  
GENERAL INFORMATION:  
APPLICANT: Lewis, Michael E.  
APPLICANT: Kauer, James C.  
APPLICANT: Smith, Kevin R.  
APPLICANT: Callison, Kathleen V.  
APPLICANT: Baldino, Frank  
APPLICANT: Neff, Nicola  
TITLE OF INVENTION: TREATING DISORDERS BY APPLICATION  
TITLE OF INVENTION: OF INSULIN-LIKE GROWTH FACTORS AND  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2604  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/958,903A  
FILING DATE: October 7, 1992  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/361,595  
FILING DATE: June 5, 1989  
APPLICATION NUMBER: 07/534,139  
FILING DATE: June 5, 1990  
APPLICATION NUMBER: April 15, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 02655/003004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-07-958-903A-51  
Query Match 48.1%; Score 25; DB 1; Length 8;  
Best Local Similarity 57.1%; Pred. No. 2e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 FCASDAK 10  
|||||  
Db 2 YCATPAK 8

RESULT 7  
US-08-462-018-51  
; Sequence 51, Application US/08462018  
; Patent No. 5703045  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, Michael E.  
; APPLICANT: Kauer, James C.  
; APPLICANT: Smith, Kevin R.  
; APPLICANT: Callison, Kathleen V.  
; APPLICANT: Baldino, Frank  
; APPLICANT: Neff, Nicola  
; APPLICANT: Iqbal, Mohamed  
; TITLE OF INVENTION: TREATING DISORDERS BY APPLICATION  
; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH FACTORS AND  
; TITLE OF INVENTION: ANALOGS  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,018  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/958,903  
; FILING DATE: October 7, 1992  
; APPLICATION NUMBER: 07/361,595  
; FILING DATE: June 5, 1989  
; APPLICATION NUMBER: 07/534,139  
; FILING DATE: June 5, 1990  
; APPLICATION NUMBER: 07/869,913  
; FILING DATE: April 15, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 02655/003005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-462-018-51  
Query Match 48.1%; Score 25; DB 1; Length 8;  
Best Local Similarity 57.1%; Pred. No. 2e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 4 FCASDAK 10  
; : : :  
; : : :  
; : : :  
Db 2 YCATPAK 8

RESULT 9  
US-07-963-329A-51  
; Sequence 51, Application US/07963329A  
; Patent No. 6310040  
; GENERAL INFORMATION:  
; APPLICANT: Bozyczko-Coyne, Donna  
; APPLICANT: Neff, Nicola  
; APPLICANT: Lewis, Michael E.  
; APPLICANT: Iqbal, Mohamed  
; TITLE OF INVENTION: TREATING DISORDERS BY  
; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH  
; TITLE OF INVENTION: FACTORS AND  
; TITLE OF INVENTION: ANALOGS  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,245  
; FILING DATE: March 24, 1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/361,595  
; FILING DATE: June 6, 1989  
; APPLICATION NUMBER: 07/534,139  
; FILING DATE: April 15, 1992  
; APPLICATION NUMBER: 07/958,903  
; FILING DATE: October 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Creeson, Gary L.  
; REGISTRATION NUMBER: 34,310  
; REFERENCE/DOCKET NUMBER: 02655/003008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8  
; TYPE: amino acid  
; STRANDEDNESS: N/A  
; TOPOLOGY: N/A  
US-08-823-245-51  
Query Match 48.1%; Score 25; DB 1; Length 8;  
Best Local Similarity 57.1%; Pred. No. 2e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 4 FCASDAK 10  
; : : :  
; : : :  
; : : :  
Db 2 YCATPAK 8

TITLE OF INVENTION: TREATING RETINAL NEURONAL DISORDERS  
TITLE OF INVENTION: BY THE APPLICATION OF INSULIN-LIKE  
TITLE OF INVENTION: GROWTH FACTORS AND ANALOGS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (version 5.0)  
SOFTWARE: WordPerfect (version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/963,329A  
FILING DATE: 19921015  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/790,690  
FILING DATE: No. 6310040ember 8, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 02655/012002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-07-963-329A-51

Query Match 48.1%; Score 25; DB 4; Length 8;  
Best Local Similarity 57.1%; Pred. No. 2e+05; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 FCASDAK 10  
Db 2 YCATPAK 8

RESULT 10  
PCT-US92-09443A-51  
Sequence 51, Application PC/TUS9209443A  
GENERAL INFORMATION:  
APPLICANT: Bozyczko-Coyne, Donna  
APPLICANT: Neft, Nicola  
APPLICANT: Lewis, Michael E.  
TITLE OF INVENTION: TREATING RETINAL NEURONAL  
TITLE OF INVENTION: DISORDERS BY THE APPLICATION OF  
TITLE OF INVENTION: INSULIN-LIKE GROWTH FACTORS AND  
TITLE OF INVENTION: ANALOGS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (version 5.0)  
SOFTWARE: WordPerfect (version 5.1)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09443A  
FILING DATE: 19921103  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/790,690  
FILING DATE: November 8, 1991  
APPLICATION NUMBER: 07/963,329  
FILING DATE: October 15, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 02655/012002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: AMINO ACID  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
PCT-US92-09443A-51

Query Match 48.1%; Score 25; DB 5; Length 8;  
Best Local Similarity 57.1%; Pred. No. 2e+05; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 FCASDAK 10  
Db 2 YCATPAK 8

RESULT 11  
US-08-866-545-19  
Sequence 19, Application US/08866545  
Patent No. 626535  
GENERAL INFORMATION:  
APPLICANT: Greene, Mark I.  
APPLICANT: Murall, Ramchandran  
APPLICANT: Takasaki, Wataru  
TITLE OF INVENTION: PEPTIDES AND PEPTIDE  
TITLE OF INVENTION: ANALOGUES DESIGNED FROM BINDING SITES OF TUMOR  
TITLE OF INVENTION: NECROSIS FACTOR RECEPTOR SUPERFAMILY AND THEIR  
TITLE OF INVENTION: USES  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/866,545  
FILING DATE: 30-MAY-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 009113-0004-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556



Db 2 YCATPAK 8

## RESULT 14

US-08-823-245-2  
; Sequence 2, Application US/08823245  
; Patent No. 5776897  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, Michael  
; APPLICANT: Kauer, James C.  
; APPLICANT: Smith, Kevin R.  
; APPLICANT: Callison, Kathleen V.  
; APPLICANT: Baldwin, Frank  
; APPLICANT: Neff, Nicola  
; APPLICANT: Iqbal, Mohamed  
; TITLE OF INVENTION: TREATING DISORDERS BY  
; TITLE OF INVENTION: APPLICATION  
; TITLE OF INVENTION: OF INSULIN-LIKE GROWTH  
; TITLE OF INVENTION: FACTORS AND  
; TITLE OF INVENTION: ANALOGS  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or  
; COMPUTER: 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823.245  
; FILING DATE: March 24, 1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/361,595  
; FILING DATE: June 6, 1989  
; APPLICATION NUMBER: 07/534,139  
; FILING DATE: June 5, 1990  
; APPLICATION NUMBER: 07/869,913  
; FILING DATE: April 15, 1992  
; APPLICATION NUMBER: 07/958,903  
; FILING DATE: October 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Creeson, Gary L.  
; REGISTRATION NUMBER: 34,310  
; REFERENCE/DOCKET NUMBER: 02655/003008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10  
; TYPE: amino acid  
; STRANDEDNESS: N/A  
; TOPOLOGY: N/A  
; US-08-823-245-2

## Query Match

Best Local Similarity 48.1%; Score 25; DB 1; Length 10;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 FCASDAK 10  
: : : : :  
Db 2 YCATPAK 8

## RESULT 15

US-07-963-329A-68  
; Sequence 68, Application US/07963329A  
; Patent No. 6310040  
; GENERAL INFORMATION:  
; APPLICANT: Bozyczko-Coyne, Donna  
; APPLICANT: Neff, Nicola  
; APPLICANT: Lewis, Michael E.  
; APPLICANT: Iqbal, Mohamed  
; TITLE OF INVENTION: TREATING RETINAL NEURONAL DISORDERS  
; TITLE OF INVENTION: BY THE APPLICATION OF INSULIN-LIKE  
; TITLE OF INVENTION: GROWTH FACTORS AND ANALOGS  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/963,329A  
; FILING DATE: 19921015  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/790,690  
; FILING DATE: No. 6310040eomber 8, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 02655/012002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEEX: 200154  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-07-963-329A-68

## Query Match

Best Local Similarity 48.1%; Score 25; DB 4; Length 10;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 FCASDAK 10  
: : : : :  
Db 2 YCATPAK 8

Search completed: April 14, 2003, 07:06:31  
Job time : 29 secs

100